

; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; TITLE OF INVENTION: TISSUE CONTRACTABILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 143  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; ORGANISM: Homo sapiens  
US-09-880-503-8

Query Match 100.0%; Score 554; DB 10; Length 143;  
Best Local Similarity 100.0%; Pred. No. 1.7e-56;  
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 48 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96  
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 143

RESULT 3  
US-10-106-698-6266  
; Sequence 6266, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 6266  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-106-698-6266

Query Match 100.0%; Score 554; DB 15; Length 337;  
Best Local Similarity 100.0%; Pred. No. 4.5e-56;  
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 74 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 133  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96  
Db 134 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 169

RESULT 4  
US-09-880-503-3  
; Sequence 3, Application US/09880503  
; Patent No. US20020131964A1

; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; TITLE OF INVENTION: TISSUE CONTRACTABILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; ORGANISM: Homo sapiens  
US-09-880-503-3

Query Match 100.0%; Score 554; DB 10; Length 411;  
Best Local Similarity 100.0%; Pred. No. 5.6e-56;  
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 48 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96  
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 143

RESULT 5  
US-09-264-468B-1  
; Sequence 1, Application US/09264468B  
; Patent No. US20020106775A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Jieyi  
; APPLICANT: Nienaber, Vicki L.  
; APPLICANT: Henkin, Jack  
; APPLICANT: Smith, Richard A.  
; APPLICANT: Walter, Karl A.  
; APPLICANT: Severin, Jean M.  
; APPLICANT: Edalji, Rohinton  
; APPLICANT: Johnson Jr., Robert W.  
; APPLICANT: Holzman, Thomas F.  
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE  
; FILE REFERENCE: 6310.US.P1  
; CURRENT APPLICATION NUMBER: US/09/264,468B  
; CURRENT FILING DATE: 1999-03-05  
; PRIOR APPLICATION NUMBER: US 09/036,361  
; PRIOR FILING DATE: 1998-03-06  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(20)  
; OTHER INFORMATION: Leader sequence  
; NAME/KEY: VARIANT  
; LOCATION: (279)...(279)  
; OTHER INFORMATION: Xaa = any amino acid  
; NAME/KEY: VARIANT  
; LOCATION: (302)...(302)  
; OTHER INFORMATION: Xaa = any amino acid  
US-09-264-468B-1

Query Match 100.0%; Score 554; DB 10; Length 431;  
Best Local Similarity 100.0%; Pred. No. 5.9e-56;  
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60  
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 127  
QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKKPSSPPEE 96  
DB 128 RRRPWCYVQVGLKPLVQECMWHDCADGKKPSSPPEE 163

RESULT 6  
US-10-301-822-161  
; Sequence 161, Application US/10301822  
; Publication No. US20030148410A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Berger, Allison  
; APPLICANT: Guillemette, Tracy L.  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John B.  
; APPLICANT: Thibodeau, Stephen N.  
; APPLICANT: Burgart, Lawrence J.  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; THERAPY OF COLON CANCER  
; FILE REFERENCE: MP01-029P2RNM  
; CURRENT APPLICATION NUMBER: US/10/301,822  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 60/339,971  
; PRIOR FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: US 60/361,978  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/381,988  
; PRIOR FILING DATE: 2002-05-20  
; NUMBER OF SEQ ID NOS: 228  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 161  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-301-822-161

Query Match 100.0%; Score 554; DB 12; Length 431;  
Best Local Similarity 100.0%; Pred. No. 5.9e-56;  
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60  
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 127  
QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKKPSSPPEE 96  
DB 128 RRRPWCYVQVGLKPLVQECMWHDCADGKKPSSPPEE 163

RESULT 7  
US-10-131-985-21  
; Sequence 21, Application US/10131985  
; Publication No. US20030199440A1  
; GENERAL INFORMATION:  
; APPLICANT: Dack, Kevin N  
; APPLICANT: Davies, Michael J  
; APPLICANT: Fish, Paul V  
; APPLICANT: Huggins, Jonathan P  
; APPLICANT: McIntosh, Fraser S  
; APPLICANT: Ocleston, Nicholas L  
; TITLE OF INVENTION: Composition  
; FILE REFERENCE: PCS 10391A  
; CURRENT APPLICATION NUMBER: US/10/131,985  
; CURRENT FILING DATE: 2002-04-25  
; PRIOR APPLICATION NUMBER: US/09/726,295  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: GB 9930768.8

; PRIOR FILING DATE: 1999-12-29  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-131-985-21

Query Match 100.0%; Score 554; DB 12; Length 431;  
Best Local Similarity 100.0%; Pred. No. 5.9e-56;  
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60  
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 127  
QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKKPSSPPEE 96  
DB 128 RRRPWCYVQVGLKPLVQECMWHDCADGKKPSSPPEE 163

RESULT 8  
US-10-076-421-2  
; Sequence 2, Application US/10076421  
; Publication No. US20020193304A1  
; GENERAL INFORMATION:  
; APPLICANT: WADA, MANABU  
; APPLICANT: WADA, NAKO  
; TITLE OF INVENTION: ANTI-HIV AGENTS  
; FILE REFERENCE: HAYAK-9  
; CURRENT APPLICATION NUMBER: US/10/076,421  
; CURRENT FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: JP 2001-42655  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: JP 2001-184284  
; PRIOR FILING DATE: 2001-06-19  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-076-421-2

Query Match 100.0%; Score 554; DB 14; Length 431;  
Best Local Similarity 100.0%; Pred. No. 5.9e-56;  
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60  
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 127  
QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGKKPSSPPEE 96  
DB 128 RRRPWCYVQVGLKPLVQECMWHDCADGKKPSSPPEE 163

RESULT 9  
US-10-171-311-184  
; Sequence 184, Application US/10171311  
; Publication No. US20030087270A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Chen, Yan  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Monahan, John  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Giatt, Karen  
; APPLICANT: Gannavarapu, Manjula  
; APPLICANT: Hoerish, Sebastian  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY

```
/ TITLE OF INVENTION: OF CERVICAL CANCER
/ FILE REFERENCE: MRI-035
/ CURRENT APPLICATION NUMBER: US/10/171,311
/ CURRENT FILING DATE: 2002-06-12
/ PRIOR APPLICATION NUMBER: US 60/298,159
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/298,155
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/335,936
/ PRIOR FILING DATE: 2001-11-14
/ NUMBER OF SEQ ID NOS: 238
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 184
/ LENGTH: 431
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-171-311-184

Query Match 100.0%; Score 554; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.9e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 163

RESULT 10
US-10-193-656-4
/ Sequence 4, Application US/10193656
/ Publication No. US20030096733A1
/ GENERAL INFORMATION:
/ APPLICANT: NY, Tor
/ APPLICANT: HOLMDAHL, Rikard
/ APPLICANT: LI, Jinan
/ TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
/ FILE REFERENCE: 3810/1J577-US3
/ CURRENT APPLICATION NUMBER: US/10/193,656
/ CURRENT FILING DATE: 2002-07-10
/ PRIOR APPLICATION NUMBER: US 60/304,461
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/304,490
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/305,182
/ PRIOR FILING DATE: 2001-07-13
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 4
/ LENGTH: 431
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: GenBank / P00749
/ DATABASE ENTRY DATE: 1986-07-21
/ RELEVANT RESIDUES: (1)..(431)
US-10-193-656-4

Query Match 100.0%; Score 554; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.9e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 163
```

```
RESULT 11
US-10-247-671-149
/ Sequence 149, Application US/10247671
/ Publication No. US20030194721A1
/ GENERAL INFORMATION:
/ APPLICANT: Mikita, Thomas
/ APPLICANT: Shiffman, Dov
/ APPLICANT: Porter, Gordon, J.
/ APPLICANT: Kaser, Matthew R.
/ TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
/ FILE REFERENCE: PA-0050 US
/ CURRENT APPLICATION NUMBER: US/10/247,671
/ CURRENT FILING DATE: 2002-09-18
/ PRIOR APPLICATION NUMBER: 60/323,784
/ PRIOR FILING DATE: 2001-09-19
/ NUMBER OF SEQ ID NOS: 186
/ SOFTWARE: PERL Program
/ SEQ ID NO 149
/ LENGTH: 431
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match 99.6%; Score 552; DB 12; Length 431;
Best Local Similarity 99.0%; Pred. No. 1e-55;
Matches 95; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 163

RESULT 12
US-10-407-821-2
/ Sequence 2, Application US/10407821
/ Publication No. US20030219386A1
/ GENERAL INFORMATION:
/ APPLICANT: IDELL, STEVEN
/ TITLE OF INVENTION: INTRALEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
/ TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
/ TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
/ FILE REFERENCE: UTSN:022US
/ CURRENT APPLICATION NUMBER: US/10/407,821
/ CURRENT FILING DATE: 2003-04-04
/ PRIOR APPLICATION NUMBER: 60/414,202
/ PRIOR FILING DATE: 2002-09-27
/ PRIOR APPLICATION NUMBER: 60/370,466
/ PRIOR FILING DATE: 2002-04-05
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 411
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-407-821-2

Query Match 98.2%; Score 544; DB 12; Length 411;
Best Local Similarity 99.0%; Pred. No. 8.1e-55;
Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
```

```
Db 108 RRRPWCYVQVGLKLLVQECMVHDCADGKKPSSPPEE 143
|||||
RESULT 13
US-09-880-503-1
; Sequence 1, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-1
Query Match 92.1%; Score 510; DB 10; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.3e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
|||||
RESULT 14
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4
Query Match 92.1%; Score 510; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 2.1e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135
|||||
Db 108 RRRPWCYVQVGLKLLVQECMVHDCADGKKPSSPPEE 143
|||||
RESULT 15
US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guittou, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES, COMPOSITION
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12
Query Match 92.1%; Score 510; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.1e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 51 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 110
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138
|||||
RESULT 16
US-10-237-667-12
; Sequence 12, Application US/10237667
; Publication No. US20030022308A1
```



GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guitton, Jean-Dominique  
Jung, Gerard  
Yeh, Patrice  
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,667  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-237-667-12  
Query Match 92.1%; Score 510; DB 15; Length 138;  
Best Local Similarity 100.0%; Pred. No. 2.1e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHYRGKASTDTMGRLPWNSTATVLQQTYYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 51 KTCYEGNGHYRGKASTDTMGRLPWNSTATVLQQTYYHAHRSDALQLGLGKHNYCRNPDN 110  
QY 61 RRPWCYVQGLKPLVQECMVHDCADGK 88  
Db 111 RRPWCYVQGLKPLVQECMVHDCADGK 138  
RESULT 17  
US-10-237-708-12  
Sequence 12, Application US/10237708  
Publication No. US20030036170A1  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guitton, Jean-Dominique  
Jung, Gerard

Yeh, Patrice  
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,708  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-237-708-12  
Query Match 92.1%; Score 510; DB 15; Length 138;  
Best Local Similarity 100.0%; Pred. No. 2.1e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHYRGKASTDTMGRLPWNSTATVLQQTYYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 51 KTCYEGNGHYRGKASTDTMGRLPWNSTATVLQQTYYHAHRSDALQLGLGKHNYCRNPDN 110  
QY 61 RRPWCYVQGLKPLVQECMVHDCADGK 88  
Db 111 RRPWCYVQGLKPLVQECMVHDCADGK 138  
RESULT 18  
US-10-237-866-12  
Sequence 12, Application US/10237866  
Publication No. US20030036171A1  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guitton, Jean-Dominique  
Jung, Gerard  
Yeh, Patrice  
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES  
NUMBER OF SEQUENCES: 36



```
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,624
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-624-12
Query Match 92.1%; Score 510; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.1e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 51 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 110

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 21
US-09-880-503-6
Sequence 6, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Rooof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
TITLE OF INVENTION: TISSUE CONTRACTABILITY
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 403
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-6
Query Match 92.1%; Score 510; DB 10; Length 403;
Best Local Similarity 100.0%; Pred. No. 7e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 107

us-09-880-503-9.rapb
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 22
US-09-987-457-18
Sequence 18, Application US/09987457
Publication No. US20030013150A1
GENERAL INFORMATION:
APPLICANT: Manosroi, Aranya
APPLICANT: Manosroi, Jiradej
APPLICANT: Tayapiwatana, Chatchai
APPLICANT: Goetz, Friedrich
APPLICANT: Werner, Rolf-Guenther
TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
FILE REFERENCE: 0652.2180001
CURRENT APPLICATION NUMBER: US/09/987,457
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/268,573
PRIOR FILING DATE: 2001-02-15
PRIOR APPLICATION NUMBER: GB 00 27 782.2
PRIOR FILING DATE: 2000-11-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 527
TYPE: PRT
ORGANISM: Homo sapiens (tPA)
US-09-987-457-18
Query Match 40.8%; Score 226; DB 11; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.4e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 61
Db 91 TCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 150

QY 62 RRRPWCYVQVGLKPLVQECMVHDCADG 87
Db 151 RRRPWCYVQVGLKPLVQECMVHDCADG 176

RESULT 23
US-09-987-455-19
Sequence 19, Application US/09987455
Publication No. US20030049729A1
GENERAL INFORMATION:
APPLICANT: Aranya Manosroi
APPLICANT: Jiradej Manosroi
APPLICANT: Chatchai Tayapiwatana
APPLICANT: Friedrich Goetz
APPLICANT: Rolf-Guenther Werner
TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
FILE REFERENCE: 0652.2190001
CURRENT APPLICATION NUMBER: US/09/987,455
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/268,574
PRIOR FILING DATE: 2001-02-15
PRIOR APPLICATION NUMBER: GB 0027779.8
PRIOR FILING DATE: 2000-11-14
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19
LENGTH: 527
TYPE: PRT
ORGANISM: Homo sapiens
US-09-987-455-19
Query Match 40.8%; Score 226; DB 11; Length 527;
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:35:43 ; Search time 21.2586 Seconds  
(without alignments)  
818.010 Million cell updates/sec

Title: US-09-880-503-3

Perfect score: 2301

Sequence: 1 SNELHQVPSNCDCLNGTGV.....VSHFLPWIRSHTKENGAL 411

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	2301	100.0	411	1	US-08-087-163-1
2	2301	100.0	411	1	US-08-286-748B-18
3	2301	100.0	411	1	US-08-153-799-18
4	2301	100.0	431	4	US-09-101-272G-1
5	2301	100.0	431	6	5188829-1
6	2298	99.9	430	1	US-07-942-157A-3
7	2291	99.6	411	3	US-09-181-816-1
8	2285.5	99.3	430	6	5219569-2
9	2277	99.0	411	2	US-08-560-098A-48
10	2023	87.9	432	2	US-08-560-098A-47
11	2022	87.9	365	1	US-08-093-741-83
12	2022	87.9	365	1	US-08-720-012-83
13	2022	87.9	393	2	US-08-560-098A-44
14	2022	87.9	393	3	US-08-967-024C-24
15	2022	87.9	393	3	US-08-967-024C-25
16	1507.5	65.5	306	2	US-08-560-098A-45
17	1507.5	65.5	331	2	US-08-560-098A-46
18	1382	60.1	253	3	US-08-944-483-73
19	1374	59.7	254	2	US-08-560-098A-49
20	1041	45.2	200	4	US-09-101-272G-73
21	883	38.4	157	3	US-08-142-590B-25
22	871.5	37.9	527	2	US-08-811-949-39
23	864.5	37.6	527	1	US-07-609-510B-16
24	864.5	37.6	527	5	PCT-US91-01025A-2
25	864.5	37.6	527	6	5185259-8
26	864.5	37.6	562	2	US-08-811-949-43
27	864.5	37.6	562	2	US-08-560-098A-50

28	864.5	37.6	562	2	US-08-883-795A-38	Sequence 38, Appl
29	864.5	37.6	562	6	5185259-3	Patent No. 5185259
30	864.5	37.6	562	6	5200340-2	Patent No. 5200340
31	864.5	37.6	562	6	5344773-2	Patent No. 5344773
32	863.5	37.5	477	2	US-08-560-098A-51	Sequence 51, Appl
33	863.5	37.5	527	6	5520913-1	Patent No. 5520913
34	857.5	37.3	562	6	5244676-5	Patent No. 5244676
35	857	37.2	208	4	US-09-101-272G-98	Sequence 98, Appl
36	804	34.9	355	2	US-08-811-949-59	Sequence 59, Appl
37	799	34.7	437	2	US-08-811-949-51	Sequence 51, Appl
38	796	34.6	437	2	US-08-811-949-57	Sequence 57, Appl
39	793	34.5	138	2	US-08-797-689-12	Sequence 12, Appl
40	792	34.4	472	2	US-08-811-949-63	Sequence 63, Appl
41	789	34.3	437	2	US-08-811-949-55	Sequence 55, Appl
42	788	34.2	194	4	US-09-101-272G-80	Sequence 80, Appl
43	788	34.2	201	4	US-09-101-272G-96	Sequence 96, Appl
44	788	34.2	437	2	US-08-811-949-49	Sequence 49, Appl
45	787	34.2	355	2	US-08-811-949-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1

US-08-087-163-1  
; Sequence 1, Application US/08087163  
; Patent No. 5472692  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jian-Ning  
; APPLICANT: Gurewich, Victor  
; TITLE OF INVENTION: PRO-UKINASE MUTANTS  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA: US/08/087.163  
; APPLICATION NUMBER: 32,983  
; FILING DATE: 07/02/93  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fasse, J. Peter  
; REGISTRATION NUMBER: 32,983  
; REFERENCE/DOCKET NUMBER: 04353/003001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 411  
; TYPE: amino acid  
; STRANDEDNESS: N/A  
; TOPOLOGY: N/A  
US-08-087-163-1

Query Match 100.0%; Score 2301; DB 1; Length 411;  
Best Local Similarity 100.0%; Pred. No. 2.1e-192;  
Matches 411; Conservative 0; Mismatches 0; Gaps 0;  
Qy 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWNCNCPKFGGQHCEIDKSKTCYEGNGHFVRG 60  
|||||

Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCCEIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSADALQGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
Db 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSADALQGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
QY 121 PLVQECMWHDCADGKKPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPMFAALYRRH 180  
Db 121 PLVQECMWHDCADGKKPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPMFAALYRRH 180  
QY 181 RGSVTVYCGGSLISPCWVISATHCFIDYPPKEDYIVYVYGRSRLNSNTGEMKFEVENLI 240  
Db 181 RGSVTVYCGGSLISPCWVISATHCFIDYPPKEDYIVYVYGRSRLNSNTGEMKFEVENLI 240  
QY 241 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIOTICLPSMYNDPQGTSCETITGFGK 300  
Db 241 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIOTICLPSMYNDPQGTSCETITGFGK 300  
QY 301 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYVGEVTTKMLCAADPOMKTDSCQDSSGGPL 360  
Db 301 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYVGEVTTKMLCAADPOMKTDSCQDSSGGPL 360  
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHRSHTEENGLAL 411  
Db 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHRSHTEENGLAL 411  
RESULT 2  
US-08-286-748B-18  
; Sequence 18, Application US/08286748B  
; Patent No. 5759542  
; GENERAL INFORMATION:  
; APPLICANT: Victor Gurewicz  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY  
; TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF  
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,748B  
; FILING DATE: August 5, 1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: J. Peter Fasse  
; REGISTRATION NUMBER: 32,983  
; REFERENCE/DOCKET NUMBER: 04547/013001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 411  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-286-748B-18  
Query Match 100.0%; Score 2301; DB 1; Length 411;

Best Local Similarity 100.0%; Pred. No. 2.1e-192;  
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCCEIDKSKTCYEGNGHFYRG 60  
Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCCEIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSADALQGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
Db 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSADALQGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
QY 121 PLVQECMWHDCADGKKPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPMFAALYRRH 180  
Db 121 PLVQECMWHDCADGKKPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPMFAALYRRH 180  
QY 181 RGSVTVYCGGSLISPCWVISATHCFIDYPPKEDYIVYVYGRSRLNSNTGEMKFEVENLI 240  
Db 181 RGSVTVYCGGSLISPCWVISATHCFIDYPPKEDYIVYVYGRSRLNSNTGEMKFEVENLI 240  
QY 241 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIOTICLPSMYNDPQGTSCETITGFGK 300  
Db 241 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIOTICLPSMYNDPQGTSCETITGFGK 300  
QY 301 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYVGEVTTKMLCAADPOMKTDSCQDSSGGPL 360  
Db 301 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYVGEVTTKMLCAADPOMKTDSCQDSSGGPL 360  
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHRSHTEENGLAL 411  
Db 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHRSHTEENGLAL 411  
RESULT 3  
US-08-153-799-18  
; Sequence 18, Application US/08153799  
; Patent No. 5766883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: R Hain Swope, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07974  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB90/00650  
; FILING DATE: 26-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/775952  
; FILING DATE: 29-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Swope, R Hain  
; REGISTRATION NUMBER: 24864  
; REFERENCE/DOCKET NUMBER: 92H832

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-153-799-18

Query Match
Best Local Similarity 100.0%; Score 2301; DB 1; Length 411;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPMCYVQVGLK 120
DB 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPMCYVQVGLK 140
QY 121 PLVQECMVHDCADGKKPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPF AAIYRRH 180
DB 141 PLVQECMVHDCADGKKPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPF AAIYRRH 200
QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGRSRLNSNTQGMKFVEVNI 240
DB 201 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGRSRLNSNTQGMKFVEVNI 260
QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSC EITGFGK 300
DB 261 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSC EITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTKMLCAADPOWKTDSC QGDSGGPL 360
DB 321 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTKMLCAADPOWKTDSC QGDSGGPL 380
QY 361 VCSLQGRMTLTGIVSWGRGCALKD KPGVYTRVSHFLPWIRSHTKENGLAL 411
DB 381 VCSLQGRMTLTGIVSWGRGCALKD KPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 5
5188829-1
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIKAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1988
; SEQ ID NO: 1;
; LENGTH: 431
5188829-1

Query Match
Best Local Similarity 100.0%; Score 2301; DB 6; Length 431;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPMCYVQVGLK 120
DB 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPMCYVQVGLK 140
QY 121 PLVQECMVHDCADGKKPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPF AAIYRRH 180
DB 141 PLVQECMVHDCADGKKPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPF AAIYRRH 200
QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGRSRLNSNTQGMKFVEVNI 240
DB 201 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGRSRLNSNTQGMKFVEVNI 260
QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSC EITGFGK 300
DB 261 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSC EITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTKMLCAADPOWKTDSC QGDSGGPL 360
DB 321 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTKMLCAADPOWKTDSC QGDSGGPL 380

US-09-101-272G-1
; SEQUENCE 1, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (21)..()
; OTHER INFORMATION:
; NAME/KEY: misc.feature
; LOCATION: (20)..()
; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
; US-09-101-272G-1

Query Match
Best Local Similarity 100.0%; Score 2301; DB 4; Length 431;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPMCYVQVGLK 120
DB 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPMCYVQVGLK 140
QY 121 PLVQECMVHDCADGKKPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPF AAIYRRH 180
DB 141 PLVQECMVHDCADGKKPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPF AAIYRRH 200
QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGRSRLNSNTQGMKFVEVNI 240
DB 201 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGRSRLNSNTQGMKFVEVNI 260
QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSC EITGFGK 300
DB 261 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPOFGTSC EITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTKMLCAADPOWKTDSC QGDSGGPL 360
DB 321 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTKMLCAADPOWKTDSC QGDSGGPL 380
```

QY 361 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTVRVSHPFWIRSHTKBENGIAL 411  
Db 381 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTVRVSHPFWIRSHTKBENGIAL 431

RESULT 6

US-07-942-157A-3  
; Sequence 3, Application US/07942157A  
; Patent No. 5648253  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Cha-Mer  
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/942,157A  
; FILING DATE: 19920908  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/631673  
; FILING DATE: 20-DEC-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: TS1108Cont.  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404)815-6508  
; TELEFAX: (404)815-6555  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 430 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..19  
; OTHER INFORMATION: /label= peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 198..203  
; OTHER INFORMATION: /label= modified  
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"  
US-07-942-157A-3

Query Match 99.9%; Score 2298; DB 1; Length 430;  
Best Local Similarity 99.8%; Pred. No. 4e-192;  
Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60  
Db 20 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 79  
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
Db 80 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 139  
QY 121 PLVQECMVHDCADGKKPSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180  
Db 140 PLVQECMVHDCADGKKPSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 199

QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 240  
Db 200 RGSVTVYVCGSLMSPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 259  
QY 241 LHKDYSADTLAHNDIALLKIRSKGRCAQPSRTTQTICLPSMYNDPQGTSCETITGFGK 300  
Db 260 LHKDYSADTLAHNDIALLKIRSKGRCAQPSRTTQTICLPSMYNDPQGTSCETITGFGK 319  
QY 301 ENSTDYLYPEOLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSDSGGPL 360  
Db 320 ENSTDYLYPEOLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSDSGGPL 379  
QY 361 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTVRVSHPFWIRSHTKBENGIAL 411  
Db 380 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTVRVSHPFWIRSHTKBENGIAL 430

RESULT 7  
US-09-181-816-1  
; Sequence 1, Application US/09181816  
; Patent No. 6277818  
; GENERAL INFORMATION:  
; APPLICANT: MAZAR, Andrew P.  
; APPLICANT: JONES, Terence R.  
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE  
; FILE REFERENCE: 32904200300 SIDN 1-7  
; CURRENT APPLICATION NUMBER: US/09/181,816  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-181-816-1

Query Match 99.6%; Score 2291; DB 3; Length 411;  
Best Local Similarity 99.8%; Pred. No. 1.5e-191;  
Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60  
Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGKKPSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180  
Db 121 LLVQECMVHDCADGKKPSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180  
QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 240  
Db 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 240  
QY 241 LHKDYSADTLAHNDIALLKIRSKGRCAQPSRTTQTICLPSMYNDPQGTSCETITGFGK 300  
Db 241 LHKDYSADTLAHNDIALLKIRSKGRCAQPSRTTQTICLPSMYNDPQGTSCETITGFGK 300  
QY 301 ENSTDYLYPEOLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSDSGGPL 360  
Db 301 ENSTDYLYPEOLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSDSGGPL 360  
QY 361 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTVRVSHPFWIRSHTKBENGIAL 411  
Db 361 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTVRVSHPFWIRSHTKBENGIAL 411

RESULT 8  
5219569-2



```

; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO:2
; LENGTH: 430
5219569-2

```

```

Query Match          99.38; Score 2285.5; DB 6; Length 430;
Best Local Similarity 99.88; Pred. No. 4.9e-191;
Matches 410; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWSNATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWSNATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140

QY 121 PLVQECMWHDCADGKPPSPPEELKFCGQKTLRPRPKIIGGFTTIENTQPFWAAIYRRH 180
DB 141 PLVQECMWHDCADG-KPPSPPEELKFCGQKTLRPRPKIIGGFTTIENTQPFWAAIYRRH 199

QY 181 RGSSTVYVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 240
DB 200 RGSSTVYVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 259

QY 241 LHKDYSADTLAHNDIALKIRKEGRCAPSPRTIQTICLPSMYNDPQFTSCIEITGFGK 300
DB 260 LHKDYSADTLAHNDIALKIRKEGRCAPSPRTIQTICLPSMYNDPQFTSCIEITGFGK 319

QY 301 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 360
DB 320 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 379

QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
DB 380 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 430

```

```

RESULT 9
US-08-560-098A-48
; Sequence 48, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins Having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A

```

```

; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-48

```

```

Query Match          99.08; Score 2277; DB 2; Length 411;
Best Local Similarity 99.38; Pred. No. 2.6e-190;
Matches 408; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWSNATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSNATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120

QY 121 PLVQECMWHDCADGKPPSPPEELKFCGQKTLRPRPKIIGGFTTIENTQPFWAAIYRRH 180
DB 121 PLVQECMWHDCADGKPPSPPEELKFCGQKTLRPRPKIIGGFTTIENTQPFWAAIYRRH 180

QY 181 RGSSTVYVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 240
DB 181 RGSSTVYVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 240

QY 241 LHKDYSADTLAHNDIALKIRKEGRCAPSPRTIQTICLPSMYNDPQFTSCIEITGFGK 300
DB 241 LHKDYSADTLAHNDIALKIRKEGRCAPSPRTIQTICLPSMYNDPQFTSCIEITGFGK 300

QY 301 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 360
DB 301 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 360

QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
DB 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

```

```

RESULT 10
US-08-560-098A-47
; Sequence 47, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins Having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```



APPLICANT: SAUNDERS, DEREK J.  
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH  
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN  
TITLE OF INVENTION: INHIBITING EFFECT  
NUMBER OF SEQUENCES: 83  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
STREET: 1200 G Street, N. W. Suite 700  
CITY: Washington, D.C.  
COUNTRY: U.S.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 27-SEP-1996  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/093,741  
FILING DATE: 20-JUL-1993  
APPLICATION NUMBER: DE P43 23 754.1  
FILING DATE: 15-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/41345  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-8800  
TELEFAX: (202) 628-8844  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-720-012-83

Query Match 87.9%; Score 2022; DB 1; Length 365;  
Best Local Similarity 100.0%; Pred. No. 3.5e-168;  
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHSDALQLGLGKHNYCRNDP 106  
Db 1 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHSDALQLGLGKHNYCRNDP 60  
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEELKFCQGQKTLRPRFKIIGGEFTT 166  
Db 61 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEELKFCQGQKTLRPRFKIIGGEFTT 120  
QY 167 TENQWFAAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNS 226  
Db 121 TENQWFAAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNS 180  
QY 227 NTQGMKEFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYND 286  
Db 181 NTQGMKEFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYND 240  
QY 287 PQFGTSCBITGFGKENSIDYLYPEQLKMTVVKLISHRECQQPHYVSEVTTKMLCAADPQ 346  
Db 241 PQFGTSCBITGFGKENSIDYLYPEQLKMTVVKLISHRECQQPHYVSEVTTKMLCAADPQ 300  
QY 347 WKTDSQCGSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEE 406  
Db 301 WKTDSQCGSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEE 360  
QY 407 NGLAL 411  
Db 361 NGLAL 365

RESULT 13

US-08-560-098A-44  
Sequence 44, Application US/08560098A  
Patent No. 5976841  
GENERAL INFORMATION:  
APPLICANT: WNEENDT, Stephan  
APPLICANT: HEINZEL-WIELAND, Regina  
APPLICANT: STEFFENS, Gerd Josef  
TITLE OF INVENTION: Proteins Having Fibrinolytic and  
TITLE OF INVENTION: Coagulation-inhibiting Properties  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
STREET: 1200 G Street, N.W., Suite 700  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/560,098A  
FILING DATE: 17-NOV-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: P 44 40 892.7  
FILING DATE: 17-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/42448  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-8800  
TELEFAX: (202) 628-8844  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 393 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-560-098A-44

Query Match 87.9%; Score 2022; DB 2; Length 393;  
Best Local Similarity 100.0%; Pred. No. 3.8e-168;  
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHSDALQLGLGKHNYCRNDP 106  
Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHSDALQLGLGKHNYCRNDP 61  
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEELKFCQGQKTLRPRFKIIGGEFTT 166  
Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEELKFCQGQKTLRPRFKIIGGEFTT 121  
QY 167 TENQWFAAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNS 226  
Db 122 TENQWFAAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNS 181  
QY 227 NTQGMKEFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYND 286  
Db 182 NTQGMKEFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYND 241  
QY 287 PQFGTSCBITGFGKENSIDYLYPEQLKMTVVKLISHRECQQPHYVSEVTTKMLCAADPQ 346  
Db 242 PQFGTSCBITGFGKENSIDYLYPEQLKMTVVKLISHRECQQPHYVSEVTTKMLCAADPQ 301  
QY 347 WKTDSQCGSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEE 406  
Db 302 WKTDSQCGSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEE 361  
QY 407 NGLAL 411

```
Db          362 NGLAL 366
|||||
347 WKTDSCQDSGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHYKEE 406
|||||
302 WKTDSCQDSGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHYKEE 361
|||||
407 NGLAL 411
|||||
362 NGLAL 366
|||||

RESULT 14
US-08-967-024C-24
; Sequence 24, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 665.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-967-024C-24

Query Match      87.9%; Score 2022; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 3.8e-168;
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 61
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPELKFQCGOKTLRPRFKIIGSEFTT 166
Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPELKFQCGOKTLRPRFKIIGSEFTT 121
QY 167 IENQWFAAAYRRHRGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNS 226
Db 122 IENQWFAAAYRRHRGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNS 181
QY 227 NTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYND 286
Db 182 NTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYND 241
QY 287 PQFGTSCIEITGFGKENSTDYLYPEOLKMTVVKLISHRECQPHYGVSEVITKMLCAADPQ 346
Db 242 PQFGTSCIEITGFGKENSTDYLYPEOLKMTVVKLISHRECQPHYGVSEVITKMLCAADPQ 301

RESULT 15
US-08-967-024C-25
; Sequence 25, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 665.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-967-024C-25

Query Match      87.9%; Score 2022; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 3.8e-168;
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 61
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPELKFQCGOKTLRPRFKIIGSEFTT 166
Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPELKFQCGOKTLRPRFKIIGSEFTT 121
QY 167 IENQWFAAAYRRHRGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNS 226
Db 122 IENQWFAAAYRRHRGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNS 181
QY 227 NTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYND 286
Db 182 NTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYND 241
QY 287 PQFGTSCIEITGFGKENSTDYLYPEOLKMTVVKLISHRECQPHYGVSEVITKMLCAADPQ 346
Db 242 PQFGTSCIEITGFGKENSTDYLYPEOLKMTVVKLISHRECQPHYGVSEVITKMLCAADPQ 301
```

QY 287 POFGTSCEITGFGKENSVDLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPQ 346  
Db 242 POFGTSCEITGFGKENSVDLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPQ 301  
QY 347 WKTDSQQGSGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHKDE 406  
Db 302 WKTDSQQGSGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHKDE 361  
QY 407 NGLAL 411  
Db 362 NGLAL 366

RESULT 16  
US-08-560-098A-45  
; Sequence 45, Application US/08560098A  
; Patent No. 5976841  
; GENERAL INFORMATION:  
; APPLICANT: WNEINDT, Stephan  
; APPLICANT: HEINZEL-WIELAND, Regina  
; APPLICANT: STEFFENS, Gerd Josef  
; TITLE OF INVENTION: Proteins having Fibrinolytic and  
; TITLE OF INVENTION: Coagulation-inhibiting Properties  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
; STREET: 1200 G Street, N.W., Suite 700  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/560,098A  
; FILING DATE: 17-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: P 44 40 892.7  
; FILING DATE: 17-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: EVANS, Joseph D.  
; REGISTRATION NUMBER: 26,269  
; REFERENCE/DOCKET NUMBER: 148/42448  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 628-8800  
; TELEFAX: (202) 628-8844  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 306 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-560-098A-45

Query Match 65.5%; Score 1507.5; DB 2; Length 306;  
Best Local Similarity 90.6%; Pred. No. 1.7e-123;  
Matches 281; Conservative 3; Mismatches 9; Indels 17; Gaps 2;

QY 103 RNPONRRP-WCYQVGLKPLVQECWHDGDKKPPSPPEELKFCQCKTLRPRFKIIG 161  
Db 13 RNPNDKYEFPW-----EDEKGFHMSPPPEELKFCQCKTLRPRFKIIG 56  
QY 162 GEFTTIENQPMFAAIYRRHRGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVLGR 221  
Db 57 GEFTTIENQPMFAAIYRRHRGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVLGR 116  
QY 222 SRLNSNTQGMKFVENLILHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIQTICLP 281

Db 117 SRLNSNTQGMKFVENLILHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIQTICLP 176  
QY 282 SMYNDQFQTSCEITGFGKENSVDLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLC 341  
Db 177 SMYNDQFQTSCEITGFGKENSVDLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLC 236  
QY 342 AADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRS 401  
Db 237 AADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRS 296  
QY 402 HTKEENGLAL 411  
Db 297 HTKEENGLAL 306

RESULT 17  
US-08-560-098A-46  
; Sequence 46, Application US/08560098A  
; Patent No. 5976841  
; GENERAL INFORMATION:  
; APPLICANT: WNEINDT, Stephan  
; APPLICANT: HEINZEL-WIELAND, Regina  
; APPLICANT: STEFFENS, Gerd Josef  
; TITLE OF INVENTION: Proteins having Fibrinolytic and  
; TITLE OF INVENTION: Coagulation-inhibiting Properties  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
; STREET: 1200 G Street, N.W., Suite 700  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/560,098A  
; FILING DATE: 17-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: P 44 40 892.7  
; FILING DATE: 17-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: EVANS, Joseph D.  
; REGISTRATION NUMBER: 26,269  
; REFERENCE/DOCKET NUMBER: 148/42448  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 628-8800  
; TELEFAX: (202) 628-8844  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 331 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-560-098A-46

Query Match 65.5%; Score 1507.5; DB 2; Length 331;  
Best Local Similarity 90.6%; Pred. No. 1.9e-123;  
Matches 281; Conservative 3; Mismatches 9; Indels 17; Gaps 2;

QY 103 RNPONRRP-WCYQVGLKPLVQECWHDGDKKPPSPPEELKFCQCKTLRPRFKIIG 161  
Db 13 RNPNDKYEFPW-----EDEKGFHMSPPPEELKFCQCKTLRPRFKIIG 56  
QY 162 GEFTTIENQPMFAAIYRRHRGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVLGR 221  
Db 57 GEFTTIENQPMFAAIYRRHRGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVLGR 116  
QY 222 SRLNSNTQGMKFVENLILHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIQTICLP 281

Db 117 SRLNSNTQGMKFEVENILHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIQTICLP 176  
QY 282 SMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIHRECCQPHYYGSEVTTKMLC 341  
Db 177 SMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIHRECCQPHYYGSEVTTKMLC 236  
QY 342 AADPQWKTDSCGSDGGLVCSLOGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRS 401  
Db 237 AADPQWKTDSCGSDGGLVCSLOGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRS 296  
QY 402 HTKEENGLAL 411  
Db 297 HTKEENGLAL 306

RESULT 18  
US-08-944-483-73  
; Sequence 73, Application US/08944483  
; Patent No. 6232456  
; GENERAL INFORMATION:  
; APPLICANT: COHEN, MAURICE  
; APPLICANT: COLPITTS, TRACEY L.  
; APPLICANT: FRIEDMAN, PAULA N.  
; APPLICANT: GRANADOS, EDWARD N.  
; APPLICANT: KLASS, MICHAEL R.  
; APPLICANT: RUSSELL, JOHN C.  
; APPLICANT: STEWART, KENT D.  
; APPLICANT: STROUPE, STEVEN D.  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
; TITLE OF INVENTION: OF THE PROSTATE  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/944,483  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Becker, Cheryl L.  
; REGISTRATION NUMBER: 35,441  
; REFERENCE/DOCKET NUMBER: 6183.US.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 847/935-1729  
; TELEFAX: 847/938-2623  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 73:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 253 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 6232456e  
US-08-944-483-73

Query Match 60.1%; Score 1382; DB 3; Length 253;  
Best Local Similarity 100.0%; Pred. No. 1-2e-112;  
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 IIGGEFTTIENQPFPAAIYRRHRGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDIYV 218  
Db 1 IIGGEFTTIENQPFPAAIYRRHRGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDIYV 60  
QY 219 LGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIQT 278  
Db 61 LGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALLKIRSKGRCAQPSRTIQT 120  
QY 279 CLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIHRECCQPHYYGSEVTTK 338  
Db 121 CLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIHRECCQPHYYGSEVTTK 180  
QY 339 MLCAADPQWKTDSCGSDGGLVCSLOGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPW 398  
Db 181 MLCAADPQWKTDSCGSDGGLVCSLOGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPW 240  
QY 399 IRSHTKEENGLAL 411  
Db 241 IRSHTKEENGLAL 253

RESULT 19  
US-08-560-098A-49  
; Sequence 49, Application US/08560098A  
; Patent No. 5976841  
; GENERAL INFORMATION:  
; APPLICANT: WNEENDT, Stephan  
; APPLICANT: HEINZEL-WIELAND, Regina  
; APPLICANT: STEFFENS, Gerd Josef  
; TITLE OF INVENTION: Proteins having Fibrinolytic and  
; TITLE OF INVENTION: Coagulation-inhibiting Properties  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
; STREET: 1200 G Street, N.W., Suite 700  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/560,098A  
; FILING DATE: 17-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: P 44 40 892.7  
; FILING DATE: 17-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: EVANS, Joseph D.  
; REGISTRATION NUMBER: 26,269  
; REFERENCE/DOCKET NUMBER: 148/42448  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 628-8800  
; TELEFAX: (202) 628-8844  
; INFORMATION FOR SEQ ID NO: 49:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 254 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-560-098A-49

Query Match 59.7%; Score 1374; DB 2; Length 254;  
Best Local Similarity 99.2%; Pred. No. 5.8e-112;  
Matches 252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 158 KIIGGEFTTIENQPFPAAIYRRHRGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDIYV 217  
Db 1 KIIGGEFTTIENQPFPAAIYRRHRGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDIYV 60

YLGSRSLNNTGEMKFEVENILHKDYSADTLAHNDIALKIRSKGRCQAPRTIQT 277  
YLGSRSLNNTGEMKFEVENILHKDYSADTLAHNDIALKIRSKGRCQAPRTIQT 120  
ICLPMSYNDPQFQTSCEITGFGKENSTLYPEQLKMTVVKLISHRECQPHYGVSEVTT 337  
ICLPMSYNDPQFQTSCEITGFGKENSTLYPEQLKMTVVKLISHRECQPHYGVSEVTT 180  
KMLCAADPQWKTSCOGDSGGLVCSLOGRMTLTGLVSGRCALKDKGVTVTRVSHFLP 397  
KMLCAADPQWKTSCOGDSGGLVCSLOGRMTLTGLVSGRCALKDKGVTVTRVSHFLP 240  
WIRSHKTEENGLAL 411  
WIRSHKTEENGLVL 254  
RESULT 20  
US-09-101-272G-73  
Sequence 73, Application US/09101272G  
Patent No. 6509445  
GENERAL INFORMATION:  
APPLICANT: Nissin Food Products Co., Ltd.  
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
FILE REFERENCE: Q50979  
CURRENT APPLICATION NUMBER: US/09/101,272G  
CURRENT FILING DATE: 1998-07-08  
PRIOR APPLICATION NUMBER: JP 1059/1996  
PRIOR FILING DATE: 1996-01-08  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 73  
LENGTH: 200  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: ATF domain of uPA  
US-09-101-272G-73  
Query Match 45.2%; Score 1041; DB 4; Length 200;  
Best Local Similarity 100.0%; Pred. No. 4.2e-83;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60  
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 80  
Qy 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
Db 81 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140  
Qy 121 PLVQECMVHDCADGKSPPEELKFCQGGOKTLRPRF 180  
Db 141 PLVQECMVHDCADGKSPPEELKFCQGGOKTLRPRF 200  
RESULT 21  
US-08-142-590B-25  
Sequence 25, Application US/08142590B  
Patent No. 6120765  
GENERAL INFORMATION:  
APPLICANT: HIBINO, Tashihiko; TAKAHASHI, Tadahito; HORII, Izumi; and GOETINCK,  
TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/142,590B  
FILING DATE: 25-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/042,318  
FILING DATE: 02-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Myers, Paul L.  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: MGP-009CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 157 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-142-590B-25  
Query Match 38.4%; Score 883; DB 3; Length 157;  
Best Local Similarity 96.8%; Pred. No. 1.8e-69;  
Matches 152; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60  
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60  
Qy 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
Db 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
Qy 121 PLVQECMVHDCADGKSPPEELKFCQGGOKTLRPRF 157  
Db 121 PLVQECMVHDCADGKSPPEELKFCQGGOKTLRPRF 157  
RESULT 22  
US-08-811-949-39  
Sequence 39, Application US/08811949  
Patent No. 5840533  
GENERAL INFORMATION:  
APPLICANT: NIWA, MINEO  
APPLICANT: SAITO, YOSHIMASA  
APPLICANT: SASAKI, HITOSHI  
APPLICANT: HAYASHI, MASAKO  
APPLICANT: NOTANI, JOUJI  
APPLICANT: KOBAYASHI, MASAKAZU  
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,949  
FILING DATE: 05-MAR-1997

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-811-949-39

Query Match 37.9%; Score 871.5; DB 2; Length 527;
Best Local Similarity 38.3%; Pred. No. 8.2e-68;
Matches 187; Conservative 56; Mismatches 156; Indels 89; Gaps 11;

QY 3 ELHVP-SNCD---CLNGGTCVSNKYFNIHWCNCPKKGGOHCEIDKSKTCYEGNGHY 58
DB 42 QCHSVFVKSCEPRFCNGTCCQALYFSD-VCQPEGFAGKCEIDTRATCYEDGIGY 100
QY 59 RGKASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVQV 118
DB 101 RGTWSTAESGAECTNWNSSALAKPYSGRRPDALRLGLGNHNYCRPNDRDRSKPWCYVKA 160
QY 119 LKPLVQECMWHDCADG-----KKPS 138
DB 161 GYSSSEFCSTPACSEGNSDCYFGNGSAYRGTHTSLTSGASCLPWNMILIGKYTTAQNPS 220
QY 139 -----SPPEELK-----FOCG-QKTLRPRFKLIIGGE 163
DB 221 AQALGKGNKYNCRNPDGDAKWPCHVLKNRRLTWEYCDVFCSTCGLRQYSPQFRILGGL 280
QY 164 FTTIENQPFATYRRH-RGGSVTVYVCGSLSPWCWISATHCFIDYPKKEIVYVLRG 222
DB 281 FADIASHPQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFPHLTVILGR 340
QY 223 RLSNQTQGMKEFVENILHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTCLPS 282
DB 341 YRVVPGEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 398
QY 283 MYNDPQFGTSCETGFGKENSVDLYPEQLKMTVVKLISHRECQPHYVGEVTTKMLCA 342
DB 399 ADQLPDWTECELSGKGHEALSPFYSERLKEAHVRLYPSRSTSQHLLNRVTIDNMLCA 458
QY 343 AD-----PQWKT-DSQCGDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 396
DB 459 GDRSGGPQANLHDAQCQDGGPLVCLNDGRMTLVGLIISWGLCGQKDVPGVYTKVTNYL 518
QY 397 PWIRSHTK 404
DB 519 DWIRDNMR 526

RESULT 23
US-07-609-510B-16
; Sequence 16, Application US/07609510B
; Patent No. 5326700
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue Pl
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,510B
; FILING DATE: 19901106
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-609-510B-16

Query Match 37.6%; Score 864.5; DB 1; Length 527;
Best Local Similarity 38.1%; Pred. No. 3.4e-67;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;

QY 3 ELHVP-SNCD---CLNGGTCVSNKYFNIHWCNCPKKGGOHCEIDKSKTCYEGNGHY 58
DB 42 QCHSVFVKSCEPRFCNGTCCQALYFSD-VCQPEGFAGKCEIDTRATCYEDGIGY 100
QY 59 RGKASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVQV 118
DB 101 RGTWSTAESGAECTNWNSSALAKPYSGRRPDALRLGLGNHNYCRPNDRDRSKPWCYVKA 160
QY 119 LKPLVQECMWHDCADG-----KKPS 138
DB 161 GYSSSEFCSTPACSEGNSDCYFGNGSAYRGTHTSLTSGASCLPWNMILIGKYTTAQNPS 220
QY 139 -----SPPEELK-----FOCG-QKTLRPRFKLIIGGE 163
DB 221 AQALGKGNKYNCRNPDGDAKWPCHVLKNRRLTWEYCDVFCSTCGLRQYSPQFRILGGL 280
QY 164 FTTIENQPFATYRRH-RGGSVTVYVCGSLSPWCWISATHCFIDYPKKEIVYVLRG 222
DB 281 FADIASHPQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFPHLTVILGR 340
QY 223 RLSNQTQGMKEFVENILHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTCLPS 282
DB 341 YRVVPGEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 398
QY 283 MYNDPQFGTSCETGFGKENSVDLYPEQLKMTVVKLISHRECQPHYVGEVTTKMLCA 342
DB 399 ADQLPDWTECELSGKGHEALSPFYSERLKEAHVRLYPSRSTSQHLLNRVTIDNMLCA 458
QY 343 AD-----PQWKT-DSQCGDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 396
DB 459 GDRSGGPQANLHDAQCQDGGPLVCLNDGRMTLVGLIISWGLCGQKDVPGVYTKVTNYL 518
QY 397 PWIRSHTK 404
DB 519 DWIRDNMR 526

RESULT 24
PCI-US91-01025A-2
; Sequence 2, Application PC/TUS9101025A
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
; TITLE OF INVENTION: Specific Properties
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

```





CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,949  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 562 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-811-949-43

Query Match	37.68;	Score	864.5;	DB 2;	Length	562;			
Best Local Similarity	38.11;	Pred.	No. 3.6e-67;						
Matches	186;	Conservative	56;	Mismatches	157;	Indels	89;	Gaps	11;
QY	3	ELHQPV-SNCD---	CLNGGTCVSNKYFSNTHWCNPKKFGQHQCEIDKSKTCYEGNGHFY	58					
Db		:		:		:		:	
QY	77	QCHSVVPKVSSEPRCFNGGTCQALYSDP-VCQCPEGFAGKCEIDTRATCYEDQGISY	135						
Db		:		:		:		:	
QY	59	RGKASTDTMGRPCLPWNSSATVLQOQTYHAHRSDALQLGLGRHNYCRNPDNRPRPWCYQVG	118						
Db		:		:		:		:	
QY	136	RGTWSTAESGAECTNMWSSALAKQPYSGRRPDAIRLGLGNHNYCRNPDROSKPWCYVFKA	195						
Db		:		:		:		:	
QY	119	LKPLVBQCMVHDCADG-----							-KKPS 138
Db		:		:		:		:	
QY	196	GKYSSEFCSTPACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNSSMILIGKVYTAQNPS	255						
Db		:		:		:		:	
QY	139	-----SPPEELK-----							-FQCG-QKTLRPREFKIIGGE 163
Db		:		:		:		:	
QY	256	AQALGLGKHNYCRNPDGAPWCHVKNRRLTWBYCDVPSCSTGELQKYSQFQFRIGKL	315						
Db		:		:		:		:	
QY	164	FTTIENQPFAAIYRRH-RGGSVTYVCGGSLISPCWVISATHCIFIDYPKKEDYTVYLGRS	222						
Db		:		:		:		:	
QY	316	FADIASHPWQAAIFAKHRRSPGFRFCGGLILSSCWILSAAHCFQERFPPHHLTVILGRT	375						
Db		:		:		:		:	
QY	223	RLNSNTGEMKFEVENILLKDYSDATLAHNNDIALLKIRSKRGRCQAPSBTITOTICLPS	282						
Db		:		:		:		:	
QY	376	YRVVPGEEQKFEYEKYIVHKEFDDDT--YDNDIALLQLKSDSSRCAQESSWRTVCLPP	433						
Db		:		:		:		:	
QY	283	MYNDPQGTSCETGFCGENSTDVLYPEQLKMTVTKLISHRECQPHYGYSEVITKMLCA	342						
Db		:		:		:		:	
QY	434	ADLQLPDWTECEUSGYKHEALSIFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCA	493						
Db		:		:		:		:	
QY	343	AD-----PQWK-DSCQDGGGLVCSLQGRMTITGTVSMGRGCALKDKPGVYTVRVSHFL	396						
Db		:		:		:		:	
QY	494	GDTRSGGPGQANLHDACQDGGGLVCLNDGEMTILVGIISWGLGCGQKDPVGVYTKVNYL	553						
Db		:		:		:		:	
QY	397	PWIRSHYK 404							
Db		:		:		:		:	
QY	554	DWIRDNNR 561							
Db		:		:		:		:	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	37.6%	Score	864.5	DB 2	Length	562			
Best Local Similarity	38.1%	Pred. No.	3.6e-67						
Matches	186	Conservative	56	Mismatches	157	Indels	89	Gaps	11
Qy	3	ELHQPV-SNCD	--CLNGGTCVSNKYFNSNTHWCNPKKFGGQHCDEIDSKTCEYGNHGFI	58					
Db	77	QCHSVFVKSCSEPRCFNGGTCQQAALP	SDP-VQCPGPFAGKCEIDTRATCYVEDQGISY	135					
Qy	59	RGKASTDTMRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNP	DNRRRPMWCYQVG	118					
Db	136	RGTWSTAESGAECTNMWSSALAKPYSGRRPDAIRLGLGNHNYCRNP	DRDRSKWCYVFKA	195					
Qy	119	LKPLVQCMTHDCADG	-----KPPS	138					
Db	196	GKYSSEFCSTPACSEGNDCYFNGSGAYRGTHSLTSESGASCLPWNSMILIGKVTYAQNPS	255						
Qy	139	-----SPPEELK	-----FOCG-QKTLRPFRKLIIGGE	163					
Db	256	AQALGLGKHNYCRNP	PGDAPKWCHVLRNRLTWECYCDVPSCSTCGLRQYVSQOPFRKGG	315					
Qy	164	FTTIENQWFAALYRRH-RGGSVTVYCGGSLISPCWVISATHCFIDYPPKKEDVIYVLGRS	222						
Db	316	FADIASHPQQAIFAKHRRSPGERFLCGGLISSCWLSAAHCQERFP	PPHLLTVILGRT	375					
Qy	223	RLMSNTQGENKFEVENILILHKVYSADTFLAHNDIALLKIRSKEGRCAQPSR	TQTCLPS	282					
Db	376	YRVVPGEEOKEFVEKIVHKEEDDOT--YDNDIALLQKSSDSSRCAOESSVVRTVCLPP	433						

QY 283 MYNDPQGTSCITGPKENSTDYLYPEQLKMTVVKLISHRECOQPHYGYSEVTTKMLCA 342  
DB 434 ADLQLPDWTCELSGKGKHEALSPFYSERLKEAHRVLYPSSRCTSQHLLNRTVTDNMLCA 493  
QY 343 AD-----POWKT-DSQCGSGGGLVCSLQGRMTLTGIVSWGRGALKDKFQVYTRYSHFL 396  
DB 494 GDTMSGPQANLHDAQCQDSGGPLVCLNDGRMTLVGLIISWGLGCGQKDVGVYTKVTNYL 553  
QY 397 PWIRSHTK 404  
DB 554 DWIRDNMR 561  
RESULT 28  
US-08-883-795A-38  
; Sequence 36, Application US/08883795A  
; Patent No. 5985607  
; GENERAL INFORMATION:  
; APPLICANT: Delcuve, Genevieve  
; APPLICANT: Awang, Gregor  
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression  
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BERSKIN & PARR  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/883,795A  
; FILING DATE: 27-JUN-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gravelle, Micheline  
; REGISTRATION NUMBER: 40,261  
; REFERENCE/DOCKET NUMBER: 7841-062  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 364-7311  
; TELEFAX: (416) 361-1398  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 562 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein (tpa)  
US-08-883-795A-38

Query Match 37.6%; Score 864.5; DB 2; Length 562;  
Best Local Similarity 38.1%; Pred. No. 3.6e-67;  
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;  
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFY 58  
DB 77 QCHSVPVKSCSBRFCNGGTCQALYFSDF-VQCPEGFAGKCEIDTRATCYVEDQG1SY 135  
QY 59 RGKASDTWGRCLPWSATVLQOYTHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVG 118  
DB 136 RGTWSTAESGAECTNWSALAKPYSGRRPDAIRLGLGNHNYCRPNDRSDKPCWCVFKA 195  
QY 119 LKPLVQECMVHDCADG-----KGPS 138  
DB 196 GKYSEFCSTPACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMSILLIKVYTAQNS 255  
QY 139 -----SPEELK-----PQCG-QKTLRPRFKIIGE 163  
DB 256 AQALGLGKHNYCRNPDPGDAKPWCHVLKNRRLTWECYCDVPSCSTCGLRQYSQPFRIKGG 315  
QY 164 FTTIENQPMFAIYRRH-RGGSVTVYVCGSLISPCWVISATHCFIDYPKEDYIVVLGRS 222  
DB 316 FADIASHPMQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFPHLTVILGRT 375  
QY 223 RLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCACQPSRTIOTICLPS 282

DB 256 AQALGLGKHNYCRNPDPGDAKPWCHVLKNRRLTWECYCDVPSCSTCGLRQYSQPFRIKGG 315  
QY 164 FTTIENQPMFAIYRRH-RGGSVTVYVCGSLISPCWVISATHCFIDYPKEDYIVVLGRS 222  
DB 316 FADIASHPMQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFPHLTVILGRT 375  
QY 223 RLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCACQPSRTIOTICLPS 282  
DB 376 YRVVPEEEQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433  
QY 283 MYNDPQGTSCITGPKENSTDYLYPEQLKMTVVKLISHRECOQPHYGYSEVTTKMLCA 342  
DB 434 ADLQLPDWTCELSGKGKHEALSPFYSERLKEAHRVLYPSSRCTSQHLLNRTVTDNMLCA 493  
QY 343 AD-----POWKT-DSQCGSGGGLVCSLQGRMTLTGIVSWGRGALKDKFQVYTRYSHFL 396  
DB 494 GDTMSGPQANLHDAQCQDSGGPLVCLNDGRMTLVGLIISWGLGCGQKDVGVYTKVTNYL 553  
QY 397 PWIRSHTK 404  
DB 554 DWIRDNMR 561  
RESULT 29  
5185259-3  
; Patent No. 5185259  
; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;  
; VEHAR, GORDON A.  
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN  
; ACTIVATOR  
; NUMBER OF SEQUENCES: 15  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/489,855  
; FILING DATE: 02-MAR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 12,694  
; FILING DATE: 09-FEB-1987  
; APPLICATION NUMBER: 483,052  
; FILING DATE: 07-APR-1983  
; APPLICATION NUMBER: 398,003  
; FILING DATE: 14-JUL-1982  
; APPLICATION NUMBER: 374,860  
; FILING DATE: 05-MAY-1982  
; SEQ ID NO: 3;  
; LENGTH: 562  
5185259-3

Query Match 37.6%; Score 864.5; DB 6; Length 562;  
Best Local Similarity 38.1%; Pred. No. 3.6e-67;  
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;  
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFY 58  
DB 77 QCHSVPVKSCSBRFCNGGTCQALYFSDF-VQCPEGFAGKCEIDTRATCYVEDQG1SY 135  
QY 59 RGKASDTWGRCLPWSATVLQOYTHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVG 118  
DB 136 RGTWSTAESGAECTNWSALAKPYSGRRPDAIRLGLGNHNYCRPNDRSDKPCWCVFKA 195  
QY 119 LKPLVQECMVHDCADG-----KGPS 138  
DB 196 GKYSEFCSTPACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMSILLIKVYTAQNS 255  
QY 139 -----SPEELK-----PQCG-QKTLRPRFKIIGE 163  
DB 256 AQALGLGKHNYCRNPDPGDAKPWCHVLKNRRLTWECYCDVPSCSTCGLRQYSQPFRIKGG 315  
QY 164 FTTIENQPMFAIYRRH-RGGSVTVYVCGSLISPCWVISATHCFIDYPKEDYIVVLGRS 222  
DB 316 FADIASHPMQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFPHLTVILGRT 375  
QY 223 RLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCACQPSRTIOTICLPS 282

Db 376 YRVVPEEEQKEVEKYIVHKEFDDDT--YNDIALQLKSDSRCAQESSVVRTVCLPP 433  
QY 283 MYNDPQFQTSCEITGFGKENSTDYLYPEQLKMTVVVKLISHRECQOPHYGSEVTTMLCA 342  
Db 434 ADLQLPDWTECELSGYKHEALSPFYSERLKEAHVRLYPSRCTSQHLLNRTVTDNMLCA 493  
QY 343 AD-----POWKT-DSCQDGGPLVCSLQGRMTLGIIVSWGRGCALKDKPGVYTVSHFL 396  
Db 494 GDTSGGPOANLHDACQDGGPLVCLNDGRMTLGIIVSWGLCGCQKDPGVYTKVTNYL 553  
QY 397 PWIRSHTK 404  
Db 554 DWIRDNMR 561

## RESULT 30

5200340-2  
; Patent No. 5200340  
; APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,  
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI  
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN  
; ACTIVATORS  
; NUMBER OF SEQUENCES: 34  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/53,412  
; FILING DATE: 22-MAY-1987  
; SEQ ID NO: 2  
; LENGTH: 562  
5200340-2

Query Match 37.6%; Score 864.5; DB 6; Length 562;  
Best Local Similarity 38.1%; Pred. No. 3.6e-67;  
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;  
QY 3 ELHQVP-SNCD---CLNGTCVSNKYFSNIHWCNPKFPGGQHCEIDKSKTCYEGNGHPY 58  
Db 77 QCHSVVPKSCSEPRFCNGTCQALYFSDF-VQCQPEGPAGKCCETDTRATCYEDQGISY 135  
QY 59 RCKASTDIMGRECLPWNATVLQOYHAHRSALQGLGKHNYCRNPDRRPPWCYVQVG 118  
Db 136 RGTWSTAESGAECTWNSSALAQKPYSGRRPDAILGLGNHNYCRNPDRDSKPCYVFKA 195  
QY 119 LKPLVQECMVHDCADG-----KKPS 138  
Db 196 GKYSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKYTAQNP 255  
QY 139 -----SPEELK-----FQCG-QKTLRPRFKLIIGE 163  
Db 256 AQALGLGKHNYCRNPDGDAKPMCHVLKNRRLTWECYDVPSCSTCGLRQYSPQFRIKGL 315  
QY 164 FTTIENQPMFAIYRRH-RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVVLGRS 222  
Db 316 FADIASHPMQAAIFAKHRSRPGERFLCGGILISSCWILSAAHCFQBRFPFHLLTVILGRT 375  
QY 223 RLNSNTQGMKEFVENLILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIQTICLPS 282  
Db 376 YRVVPEEEQKEVEKYIVHKEFDDDT--YNDIALQLKSDSRCAQESSVVRTVCLPP 433  
QY 283 MYNDPQFQTSCEITGFGKENSTDYLYPEQLKMTVVVKLISHRECQOPHYGSEVTTMLCA 342  
Db 434 ADLQLPDWTECELSGYKHEALSPFYSERLKEAHVRLYPSRCTSQHLLNRTVTDNMLCA 493  
QY 343 AD-----POWKT-DSCQDGGPLVCSLQGRMTLGIIVSWGRGCALKDKPGVYTVSHFL 396  
Db 494 GDTSGGPOANLHDACQDGGPLVCLNDGRMTLGIIVSWGLCGCQKDPGVYTKVTNYL 553  
QY 397 PWIRSHTK 404  
Db 554 DWIRDNMR 561

Search completed: December 3, 2003, 14:45:34  
Job time : 23.2586 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:02 ; Search time 64.7422 Seconds  
(without alignments)  
1007.637 Million cell updates/sec

Title: US-09-880-503-3  
Perfect score: 2301  
Sequence: 1 SNEHQVPSNCCLNGTCV.....VSHFLPWIRSHKTEGLAL 411

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2301	100.0	411	6 AAP50871	Sequence encoded b
2	2301	100.0	411	11 AAR06244	Urokinase precursor
3	2301	100.0	411	16 AAR62991	Pro-urokinase. Ho
4	2301	100.0	411	21 AAY92836	Urokinase plasmino
5	2301	100.0	411	23 AAE16544	Human urokinase-ty
6	2301	100.0	431	6 AAP50114	Sequence encoded b
7	2301	100.0	431	7 AAE60783	Human urokinase.
8	2301	100.0	431	7 AAE70258	Sequence of human
9	2301	100.0	431	9 AAP80430	Deduced AA sequenc

10	2301	100.0	431	9 AAP81204	Pro-urokinase with
11	2301	100.0	431	10 AAP92119	Natural human prou
12	2301	100.0	431	11 AAR04253	Human pro-urokinas
13	2301	100.0	431	11 AAR07112	Human pro-Urokinas
14	2301	100.0	431	15 AAR47903	Pro-urokinase deri
15	2301	100.0	431	15 AAR63141	Full length human
16	2301	100.0	431	21 AAY95991	Human plasminogen
17	2301	100.0	431	21 AAY50869	Human urokinase pr
18	2301	100.0	431	22 AAB84605	Amino acid sequenc
19	2301	100.0	431	23 AAG79460	sc-UPA. Homo sapi
20	2301	100.0	431	23 AAU99228	Human plasminogen
21	2301	100.0	431	23 ABE17128	Human uPA protein.
22	2301	100.0	431	24 ABU56547	Lung cancer-associ
23	2301	100.0	431	24 ABU56708	Lung cancer-associ
24	2301	100.0	431	24 ABU11076	Human urokinase pl
25	2301	100.0	434	13 AAR20537	Amidated deriv. of
26	2301	100.0	434	13 AAR20538	Amidated deriv. of
27	2301	100.0	436	13 AAR20536	Amidated deriv. of
28	2298	99.9	430	18 AAW24578	Inhibitor resistan
29	2298	99.9	431	10 AAP91886	Sequence of prouro
30	2298	99.9	431	10 AAP94764	Non-glycosylated p
31	2297	99.8	411	14 AAR34584	Mutant human prou
32	2297	99.8	431	23 AAU99230	Human plasminogen
33	2296	99.8	411	17 AAP96146	Sequence encoded b
34	2296	99.8	411	17 AAR92526	Pro-urokinase. Ho
35	2296	99.8	431	22 AAB74797	Prourokinase prote
36	2296	99.8	431	23 AAU99229	Human plasminogen
37	2295	99.7	411	16 AAR62992	Pro-urokinase muta
38	2295	99.7	411	16 AAR62993	Pro-urokinase muta
39	2295	99.7	411	16 AAR62998	Pro-urokinase muta
40	2295	99.7	411	16 AAR62999	Pro-urokinase muta
41	2293	99.7	411	16 AAR63000	Pro-urokinase muta
42	2293	99.7	411	20 AAY39343	Human pro-urokinas
43	2293	99.7	411	20 AAY42284	Human pro-urokinas
44	2293	99.7	411	22 AAB20489	Human pro-urokinas
45	2293	99.7	412	12 AAR10334	Recombinant single

ALIGNMENTS

RESULT 1  
AAP50871 ID AAP50871 standard; protein; 411 AA.  
XX AAP50871; AC AC  
XX 30-NOV-1991 (first entry)  
XX DT  
DE Sequence encoded by cDNA sequence for human urokinase zymogen  
DE (Japanese Patent Application No.37119/84).  
XX KW Thrombolytic agent; plasminogen activator activity; fibrin affinity;  
XX enzyme.  
XX OS Homo sapiens.  
XX FH Key  
FT Cleavage-site 158..159  
FT /note= "potential cleavage site which generates  
FT the two-chain form from the zymogen"

Disulfide-bond 50..131  
Disulfide-bond 71..113  
Disulfide-bond 102..126  
Disulfide-bond 148..279  
Disulfide-bond 189..205  
Disulfide-bond 197..268  
Disulfide-bond 293..362  
Disulfide-bond 325..341  
Disulfide-bond 352..380  
EP139447-A.

PD 02-MAY-1985.  
XX  
PF 07-SEP-1984; 84EP-0306117.  
XX  
PR 17-OCT-1983; 83JP-0195051.  
PR 13-SEP-1983; 83JP-0170354.  
XX  
PA (GREC ) GREEN CROSS CORP.  
XX  
PI Kasai S, Arimura H, Mori K, Nishida M, Suyama T;  
XX WPI; 1985-106530/18.  
DR  
XX New urokinase zymogen - useful as thrombolytic agent  
PT  
XX Disclosure; Page 12; 30pp; English.  
PS  
XX Zymogen AAP50871 is the inactive precursor form of human urokinase.  
CC Urokinase zymogen is cleaved into the two-chain form composed of  
CC characteristic urokinase H (molecular wt. of 30,000) and L (molecular  
CC wt. of 20,000) chains when treated with catalytic amounts of plasmin.  
CC The patentors claim a new urokinase zymogen which has mol. wt. ca.  
CC 50,000, a single chain molecular structure, and selective affinity  
CC for fibrin. It is a thrombolytic agent which manifests its  
CC plasminogen activator activity on cleavage by proteolytic enzymes  
CC (e.g. plasmin) and has higher affinity for fibrin than known forms  
CC of urokinase.  
XX  
SQ Sequence 411 AA;  
Query Match 100.0%; Score 2301; DB 6; Length 411;  
Best Local Similarity 100.0%; Pred. No. 5.4e-178;  
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGPRCLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVOVGLK 120  
DB 61 KASTDTMGPRCLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVOVGLK 120  
QY 121 PLVQECMWHDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180  
DB 121 PLVQECMWHDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180  
QY 181 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNNTQGMKFVEVENLI 240  
DB 181 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNNTQGMKFVEVENLI 240  
QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAPRTIQTICLPSMYNDPQGTSCETGFGK 300  
DB 241 LHKDYSADTLAHNDIALKIRSKGRCQAPRTIQTICLPSMYNDPQGTSCETGFGK 300  
QY 301 ENSTDYLYPEQLKMTVVVLIHRECCQPHYGVSEVTTKMLCAADPOWKTDSCQDGGGGL 360  
DB 301 ENSTDYLYPEQLKMTVVVLIHRECCQPHYGVSEVTTKMLCAADPOWKTDSCQDGGGGL 360  
QY 361 VCSLQGRMTLTGIVSWGRGCALDKPKGVTVRVSHFLPWIRSHTKENGIAL 411  
DB 361 VCSLQGRMTLTGIVSWGRGCALDKPKGVTVRVSHFLPWIRSHTKENGIAL 411  
RESULT 2  
AAR06244  
ID AAR06244 standard; protein; 411 AA.  
XX AAR06244;  
XX  
XX 07-DEC-1990 (first entry)  
DT  
XX Urokinase precursor protein.  
DE  
XX

KU Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;  
KW myocardial infarction;  
XX Homo sapiens.  
OS  
XX EP380334-A.  
PN  
PD 01-AUG-1990.  
XX  
PF 25-JAN-1990; 90EP-0300772.  
XX  
PR 17-MAY-1989; 89JP-0121405.  
PR 27-JAN-1989; 89JP-0016406.  
XX  
PA (GREC ) GREEN CROSS CORP.  
XX  
PI Matsuda H, Ueda Y, Tamanouchi K;  
XX WPI; 1990-233117/31.  
DR  
XX Urokinase precursor-lipid composite - used as thrombolytic agent,  
PT having prolonged half-life in the blood, enhanced  
PT bioavailability and improved activity  
XX  
PS Claim 3; Fig 1; lipp; English.  
XX  
CC By forming a precursor-lipid composite, the half-life of this  
CC thrombolytic agent in the blood may be increased, exhibiting  
CC improved activity without abnormal acceleration of fibrinolytic  
CC activity. Compound is useful as a thrombolytic agent in  
CC treatment of cerebral thrombosis, myocardial infarction etc.  
XX  
SQ Sequence 411 AA;  
Query Match 100.0%; Score 2301; DB 11; Length 411;  
Best Local Similarity 100.0%; Pred. No. 5.4e-178;  
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGPRCLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVOVGLK 120  
DB 61 KASTDTMGPRCLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVOVGLK 120  
QY 121 PLVQECMWHDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180  
DB 121 PLVQECMWHDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180  
QY 181 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNNTQGMKFVEVENLI 240  
DB 181 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNNTQGMKFVEVENLI 240  
QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAPRTIQTICLPSMYNDPQGTSCETGFGK 300  
DB 241 LHKDYSADTLAHNDIALKIRSKGRCQAPRTIQTICLPSMYNDPQGTSCETGFGK 300  
QY 301 ENSTDYLYPEQLKMTVVVLIHRECCQPHYGVSEVTTKMLCAADPOWKTDSCQDGGGGL 360  
DB 301 ENSTDYLYPEQLKMTVVVLIHRECCQPHYGVSEVTTKMLCAADPOWKTDSCQDGGGGL 360  
QY 361 VCSLQGRMTLTGIVSWGRGCALDKPKGVTVRVSHFLPWIRSHTKENGIAL 411  
DB 361 VCSLQGRMTLTGIVSWGRGCALDKPKGVTVRVSHFLPWIRSHTKENGIAL 411  
RESULT 3  
AAR62991  
ID AAR62991 standard; protein; 411 AA.  
XX  
AC AAR62991;  
XX

DT 25-MAR-2003 (updated)  
 XX 21-SEP-1995 (first entry)  
 DE Pro-urokinase.  
 KW Pro-urokinase; thrombolysis; fibrin clot lysis.  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 297..313  
 FT Disulfide-bond 11..19 /note= "flexible loop"  
 FT Disulfide-bond 13..31  
 FT Disulfide-bond 33..42  
 FT Disulfide-bond 50..131  
 FT Disulfide-bond 71..113  
 FT Disulfide-bond 102..126  
 FT Disulfide-bond 148..279  
 FT Disulfide-bond 189..205  
 FT Disulfide-bond 197..268  
 FT Disulfide-bond 293..362  
 FT Disulfide-bond 325..341  
 FT Disulfide-bond 352..380  
 XX  
 PN W09501427-A1.  
 XX  
 XX  
 PD 12-JAN-1995.  
 XX  
 XX  
 PF 28-JUN-1994; 94WO-US07278.  
 XX  
 XX  
 PR 02-JUL-1993; 93US-0087163.  
 XX  
 XX  
 PA (NEW-) NEW ENGLAND DEACONESS HOSPITAL.  
 XX  
 PI Gurewich V, Liu J;  
 XX  
 DR WPI; 1995-060991/08.  
 XX  
 XX  
 PT Pro-urokinase mutants - have thrombolytic activity but reduced  
 PT fibrinogenolysis activity and non-specific plasminogen activation  
 XX  
 PS Disclosure; Fig 1; 46pp; English.  
 XX  
 CC AARG2991 is the wild type pro-urokinase, from which the new mutants  
 CC described in AARG2992-R6308 were derived. These mutants retain the  
 CC thrombolytic activity of the wild type protein, useful for the  
 CC treatment of thromboembolism, but have a reduced fibrinogenolysis  
 CC activity and non-specific plasminogen activation. The mutants can  
 CC therefore be used for the lysis of fibrin clots without inducing  
 CC systemic bleeding, as can be the case with the wild type protein.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 411 AA;  
 Query Match 100.0%; Score 2301; DB 16; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 5,4e-178;  
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNELHVFNSCDLNGTCTVSNKYFNSNTHWCNCPKFGQCHCEIDKSKTCYEGNGHFYRG 60  
 DB 1 SNELHVFNSCDLNGTCTVSNKYFNSNTHWCNCPKFGQCHCEIDKSKTCYEGNGHFYRG 60  
 QY 61 KASTDTMGRCPLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
 DB 61 KASTDTMGRCPLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
 QY 121 PLVQCMWHDCADGKKPSPPEELKFCQCKTLRPRFKLIIGFEFTTIENQPFWFAIYRRH 180  
 DB 121 PLVQCMWHDCADGKKPSPPEELKFCQCKTLRPRFKLIIGFEFTTIENQPFWFAIYRRH 180  
 QY 181 RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTGEMKFEVENLI 240  
 XX

DB 181 RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTGEMKFEVENLI 240  
 QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTTQTICLPSMYNDPQGTSCETGFGK 300  
 DB 241 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTTQTICLPSMYNDPQGTSCETGFGK 300  
 QY 301 ENSTDYLYPEQLKMTWVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPL 360  
 DB 301 ENSTDYLYPEQLKMTWVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPL 360  
 QY 361 VCSLQGRMTLTGIVSWGRCALKDKPVGVTYRVSHFLPWIRSHTKENGLAL 411  
 DB 361 VCSLQGRMTLTGIVSWGRCALKDKPVGVTYRVSHFLPWIRSHTKENGLAL 411  
 RESULT 4  
 AAY92836  
 ID AAY92836 standard; Protein; 411 AA.  
 XX  
 AC AAY92836;  
 XX  
 DT 29-AUG-2000 (first entry)  
 XX  
 DE Urokinase plasminogen activator (uPA).  
 XX  
 KW N-terminal; pro-uPA; urokinase plasminogen activator; receptor; uPAR;  
 KW anti-cancer; anti-metastatic; anti-proliferative; anti-atherosclerotic;  
 KW anti-thrombotic; anti-angiogenic; anti-inflammatory; anti-arthritis;  
 KW anti-fibrotic; apoptotic; vasotropic; anti-diabetic; ophthalmological;  
 KW thrombolytic.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Disulfide-bond 11..19  
 FT Disulfide-bond 13..31  
 FT Disulfide-bond 33..42  
 FT Disulfide-bond 50..131  
 FT Disulfide-bond 71..113  
 FT Disulfide-bond 102..126  
 FT Disulfide-bond 148..279  
 FT Disulfide-bond 189..205  
 FT Disulfide-bond 197..268  
 FT Disulfide-bond 293..362  
 FT Disulfide-bond 325..341  
 FT Disulfide-bond 352..380  
 XX  
 PN W0200026353-A1.  
 XX  
 XX  
 PD 11-MAY-2000.  
 XX  
 PF 28-OCT-1999; 99WO-US25210.  
 XX  
 XX 29-OCT-1998; 98US-0181816.  
 PR (ANGS-) ANGSTROM PHARM INC.  
 XX  
 PI Mazar AP, Jones TR;  
 XX  
 DR WPI; 2000-365605/31.  
 XX  
 XX  
 PT New cyclic peptide, useful for treatment or diagnosis of e.g. tumors  
 PT and other diseases involving cell proliferation or migration, targets  
 PT the urokinase plasminogen activator receptor  
 XX  
 PS Disclosure; Fig 1; 93pp; English.  
 XX  
 CC The present sequence shows the wild-type urokinase plasminogen activator  
 CC (uPA). Cyclic peptides based on the amino acids residues 20-30 (the  
 CC receptor-binding region) of uPA are claimed. These cyclic peptides target  
 CC the uPA receptor (uPAR), allowing therapeutic or diagnostic agents to be  
 CC delivered to uPAR-expressing cells. The cyclic peptides are used,  
 CC optionally when linked to a therapeutic agent, to inhibit migration,

CC invasion and proliferation of cells, or angiogenesis, or to induce  
CC apoptosis. Particularly they are used, in human or veterinary medicine,  
CC to treat diseases characterized by these processes, e.g. solid tumors,  
CC leukemia or lymphoma (or their metastases); benign hyperplasia;  
CC atherosclerosis; restenosis; ischaemia; deep vein thrombosis; neovascular  
CC glaucoma; diabetic retinopathy; arthritis; fibrosis; bone fracture etc.,  
CC most particularly growth, invasion and metastasis of tumors. When  
CC labeled, the cyclic peptides can be used for diagnostic detection of uPAR  
CC (a marker of metastasis) on cells, tissues etc., in vivo or in vitro, and  
CC when immobilized they are used to isolate uPAR or cells that express  
CC them. The cyclic peptides are stable, soluble in water, bind strongly to  
CC uPAR, are relatively inexpensive to produce and may be derivatized by  
CC attachment of therapeutic or diagnostic agents without significantly  
CC affecting their binding. Since they target uPAR, they should have  
CC relatively low systemic toxicity and only low doses are required.

XX Sequence 411 AA;

Query Match 100.0%; Score 2301; DB 21; Length 411;  
Best Local Similarity 100.0%; Pred. No. 5.4e-178;  
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGTCTVSNKYFNSNIHWCNCPKPGGQHCHIDSKTKCYEGNGHYRG 60  
DB |||||  
1 SNEHQVPSNCDCLNGTCTVSNKYFNSNIHWCNCPKPGGQHCHIDSKTKCYEGNGHYRG 60  
QY 61 KASTDTMGRPCLPWNATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRPWCYVQVGLK 120  
DB |||||  
61 KASTDTMGRPCLPWNATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGKKPSSPEELKFCQCKTLRPRFKLIIGGFTTIENQPFALYYRH 180  
DB |||||  
121 PLVQECMVHDCADGKKPSSPEELKFCQCKTLRPRFKLIIGGFTTIENQPFALYYRH 180  
QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 240  
DB |||||  
181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 240  
QY 241 LHKDYSADTLAHHNDIALKIRSKGRCQAPSTIQTICLPSMYNDPQFGTSCITGFGK 300  
DB |||||  
241 LHKDYSADTLAHHNDIALKIRSKGRCQAPSTIQTICLPSMYNDPQFGTSCITGFGK 300  
QY 301 ENSTDYLYPEQLKMTVVVKLIISHRECCQPHYGYSEVTTKMLCAADPQWKTDSCQDSSGGPL 360  
DB |||||  
301 ENSTDYLYPEQLKMTVVVKLIISHRECCQPHYGYSEVTTKMLCAADPQWKTDSCQDSSGGPL 360  
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411  
DB |||||  
361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 5  
AAE16544

ID AAE16544 standard; Protein; 411 AA.

XX AAE16544;

AC AAE16544;

DT 09-APR-2002 (first entry)

DE Human urokinase-type plasminogen activator tcuPA and scuPA protein.

KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
KW clotting disorder; uterine contraction disorder; respiratory disease;  
KW male impotence; adult respiratory distress syndrome; tcuPA; scuPA;  
KW two chain urokinase; single chain urokinase.

OS Homo sapiens.

XX

PN WO200197752-A2.

XX

PD 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US18976.

XX 20-JUN-2000; 2000US-212874P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Cines DB, Higazi AA;

XX WPI; 2002-122240/16.

XX N-PSDB; AAD27077.

CC Composition for modulating muscle cell and tissue contractility for  
CC treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
CC comprising domains from urokinase-type plasminogen activator -  
CC Claim 9; Fig 1C; 117pp; English.

XX The invention relates to a composition comprising one or more domains of  
XX urokinase-type plasminogen activator (uPA). The composition is used to  
XX modulate the contractility and angiogenic activity of a mammalian muscle,  
XX endothelial cell or tissue. The composition is used for treating stroke,  
XX hypotension, hypertension, atherosclerosis, heart attack, microvascular  
XX occlusions, thrombotic microangiopathies, surgically induced thrombotic  
XX disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
XX invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
XX diabetic retinopathy, wound healing, clotting disorder, uterine  
XX contraction disorder, male impotence, respiratory disease or condition  
XX such as asthma, adult respiratory distress syndrome, primary pulmonary  
XX hypertension, microvascular thrombotic occlusion, and a disorder  
XX associated with chronic intrapulmonary fibrin formation. The present  
XX sequence is human urokinase-type plasminogen activator (uPA) two  
XX chain urokinase (tcuPA) and single chain urokinase (scuPA) protein.

XX Sequence 411 AA;

Query Match 100.0%; Score 2301; DB 23; Length 411;  
Best Local Similarity 100.0%; Pred. No. 5.4e-178;  
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGTCTVSNKYFNSNIHWCNCPKPGGQHCHIDSKTKCYEGNGHYRG 60

DB |||||

1 SNEHQVPSNCDCLNGTCTVSNKYFNSNIHWCNCPKPGGQHCHIDSKTKCYEGNGHYRG 60

QY 61 KASTDTMGRPCLPWNATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRPWCYVQVGLK 120

DB |||||

61 KASTDTMGRPCLPWNATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRPWCYVQVGLK 120

QY 121 PLVQECMVHDCADGKKPSSPEELKFCQCKTLRPRFKLIIGGFTTIENQPFALYYRH 180

DB |||||

121 PLVQECMVHDCADGKKPSSPEELKFCQCKTLRPRFKLIIGGFTTIENQPFALYYRH 180

QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 240

DB |||||

181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 240

QY 241 LHKDYSADTLAHHNDIALKIRSKGRCQAPSTIQTICLPSMYNDPQFGTSCITGFGK 300

DB |||||

241 LHKDYSADTLAHHNDIALKIRSKGRCQAPSTIQTICLPSMYNDPQFGTSCITGFGK 300

QY 301 ENSTDYLYPEQLKMTVVVKLIISHRECCQPHYGYSEVTTKMLCAADPQWKTDSCQDSSGGPL 360

DB |||||

301 ENSTDYLYPEQLKMTVVVKLIISHRECCQPHYGYSEVTTKMLCAADPQWKTDSCQDSSGGPL 360

QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

DB |||||

361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 6  
AAP50114

ID AAP50114 standard; Protein; 431 AA.





Db 141 PLVQECWVHDCADGKKPSSPPBELKFCQCGQKTLRPRFKIIGGFTTIENQWPFAAIYRRH 200  
Qy 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVEVENLI 240  
Db 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVEVENLI 260  
Qy 241 LHKDYSADTLAHHNDIALKIRSKGRCQAPSTIQTICLPSMYNDPQFGTSCETITGPGK 300  
Db 261 LHKDYSADTLAHHNDIALKIRSKGRCQAPSTIQTICLPSMYNDPQFGTSCETITGPGK 320  
Qy 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVGGSEVTTKMLCAADPOWKTDSCQDGGPL 360  
Db 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVGGSEVTTKMLCAADPOWKTDSCQDGGPL 380  
Qy 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411  
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 8  
AAP70258  
ID AAP70258 standard; protein; 431 AA.

XX AAP70258;  
AC  
XX  
DT 25-MAR-2003 (updated)  
DT 19-MAY-1991 (first entry)

DE Sequence of human prourokinase and leader.  
XX  
KW Cardiovascular disease treatment; fibrin affinity; thrombolytic;  
KW enzyme; protease.  
XX  
OS Homo sapiens.

XX Key  
FH Peptide  
FT /label= leader  
FT 21..431  
FT /label= prourokinase

XX EP231883-A.  
XX  
XX 12-AUG-1987.  
XX  
XX 29-JAN-1987; 87EP-0101209.  
XX  
XX 31-JAN-1986; 86JP-0017734.  
XX 30-JAN-1987; 87JP-0018626.

XX (SAGA ) SAGAMI CHEM RES CENTRE.  
XX (NIPS ) NIPPON SODA CO.  
XX (CENG ) CENTRAL GLASS CO LTD.  
XX (TOXU ) TOYO SODA MFG CO LTD.  
XX (NISC ) NISSAN CHEM IND LTD.  
XX (NISC ) NISSAN CHEMICAL INDS KK.

XX Tagawa M, Wada M, Yamada M, Yokoyama M, Numao N;  
XX WPI; 1987-222882/32.  
XX N-PSDB; AAN70390.

XX Hybrid plasminogen activator-like polypeptide - having a region  
XX for affinity to fibrin from tissue plasminogen activator and a  
XX region from prourokinase

XX Disclosure; Fig 2(1-5); 64pp; English.

XX The TPA portion of the claimed hybrid polypeptide (see FT) may  
XX consist of 2 kringles from N-terminal first serine to 219th glycine  
XX of human TPA, 1 kringle from 128th serine to 219th glycine of human  
XX TPA or half a kringle from 161st methionine to 219th glycine (see

CC AAP70257). The C-terminal half of the hybrid polypeptide may contain  
CC an AA sequence from 150th glutamine to C-terminal 411th leucine of  
CC prourokinase (see AAP70258).  
CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 431 AA;

Qy Query Match 100.0%; Score 2301; DB 8; Length 431;  
Best Local Similarity 100.0%; Pred. No. 5.7e-178;  
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHGVPCNCLNGGTCSNKNYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHYRG 60  
Db 21 SNELHGVPCNCLNGGTCSNKNYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHYRG 80

Qy 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKENVYCRNPNRRRPMWCYVQVGLK 120  
Db 81 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKENVYCRNPNRRRPMWCYVQVGLK 140

Qy 121 PLVQECWVHDCADGKKPSSPPBELKFCQCGQKTLRPRFKIIGGFTTIENQWPFAAIYRRH 180  
Db 141 PLVQECWVHDCADGKKPSSPPBELKFCQCGQKTLRPRFKIIGGFTTIENQWPFAAIYRRH 200

Qy 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVEVENLI 240  
Db 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVEVENLI 260

Qy 241 LHKDYSADTLAHHNDIALKIRSKGRCQAPSTIQTICLPSMYNDPQFGTSCETITGFGK 300  
Db 261 LHKDYSADTLAHHNDIALKIRSKGRCQAPSTIQTICLPSMYNDPQFGTSCETITGFGK 320

Qy 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVGGSEVTTKMLCAADPOWKTDSCQDGGPL 360  
Db 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVGGSEVTTKMLCAADPOWKTDSCQDGGPL 380

Qy 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411  
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 9  
AAP80430  
ID AAP80430 standard; protein; 431 AA.

XX AAP80430;  
AC  
XX  
DT 25-MAR-2003 (updated)  
DT 14-SEP-1990 (first entry)

DE Deduced AA sequence of the single chain urokinase plasminogen activator  
DE (SCU-PA) cDNA insert prepared from human Hep3 cells.  
XX  
XX Single chain urokinase plasminogen activator (SCU-PA); human Hep3 cells;  
XX glycerolaldehyde-3-phosphate dehydrogenase gene; thrombosis prevention;  
XX thrombosis treatment.

XX Homo sapiens.

XX Key  
FH Key Location/Qualifiers  
FT Peptide 1..20  
FT Protein 21..411

XX EP288435-A.

XX 26-OCT-1988.

XX 11-APR-1988; 88EP-0810234.

XX 15-APR-1987; 87GB-0009081.

XX 16-JUN-1987; 87GB-0014059.

XX 04-DEC-1987; 87IE-0003299.

XX (CIBA ) CIBA GEIGY AG.

XX PI Meyhack B, Heim J, Burgi R;  
 XX DR WPI; 1988-301440/43.  
 XX DR N-PSDB; AAN80981.  
 XX PT Prodn. of human single chain urokinase type plasminogen activator -  
 XX PT by culturing yeast strain transformed with hybrid vector contg.  
 XX PT yeast expression control sequences  
 XX PS Example 1; Fig 2; 48pp; English.  
 XX CC The patent is for the prodn. of human single chain urokinase-type  
 CC plasminogen activator (UTPA). Mutants of scu-PA are especially those  
 CC which render the protein protease resistant. Such scu-PA mutants are  
 CC covalently modified at sites of proteolysis by proteases occurring in  
 CC blood such as thrombin or plasmin, so that they are no longer  
 CC susceptible to protease hydrolysis at these locations. The target sites  
 CC include Lys135 to Lys136 (cleavage at this site generates the so-called  
 CC low molecular weight form of scu-PA or LUK), Arg156 to Phe157  
 CC (susceptible to thrombin attack) and Lys158 to Ile159 (cleavage at this  
 CC site by plasmin generates tuc-PA). Suitable scu-PA mutants have site  
 CC specific substitutions, insertions or deletions of residues at  
 CC one or more of these target sites. Especially preferred are those  
 CC mutants in which one amino acid residue or both amino acid residues  
 CC forming the target sites are deleted or in which at least one of these  
 CC amino acid residues is replaced by another amino acid residue so that  
 CC the resulting mutants are resistant to proteolytic attack. The UTPA  
 CC proteins exhibit the biological activity of natural human UTPA without  
 CC any refolding procedure being necessary. They can be used as for known  
 CC PAs in humans for the prevention or treatment of thrombosis or other  
 CC conditions where it is desired to produce local fibrinolytic or  
 CC proteolytic activity.  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 XX CC  
 XX SQ Sequence 431 AA;  
 Query Match 100.0%; Score 2301; DB 9; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-178;  
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFNIHWNCNPKFGQHCIDKSKTCYEGNGHFYRG 60  
 DB 21 SNEHQVPSNCDCLNGGTCVSNKYFNIHWNCNPKFGQHCIDKSKTCYEGNGHFYRG 80  
 QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
 DB 81 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 140  
 QY 121 PLVQECMWHDCADGKKPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPF AAIYRRH 180  
 DB 141 PLVQECMWHDCADGKKPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPF AAIYRRH 200  
 QY 181 RGSVTVYCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 240  
 DB 201 RGSVTVYCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 260  
 QY 241 LHKDYSADTLAHHNDIALKIRSKGRCAQPSRTIQTICLPMSYNDPQFSGTCEITGFGK 300  
 DB 261 LHKDYSADTLAHHNDIALKIRSKGRCAQPSRTIQTICLPMSYNDPQFSGTCEITGFGK 320  
 QY 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDCQGDGSGPL 360  
 DB 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDCQGDGSGPL 380  
 QY 361 VCSLOGRMTLTGIVSGRCALKDKPGVYTRVSHFLPWIRSRHTKEENGIAL 411  
 DB 381 VCSLOGRMTLTGIVSGRCALKDKPGVYTRVSHFLPWIRSRHTKEENGIAL 431  
 RESULT 10  
 AAP81204  
 ID AAP81204 standard; protein; 431 AA.

XX AAP81204;  
 XX AC  
 XX DT 25-MAR-2003 (updated)  
 XX DT 03-DEC-1990 (first entry)  
 XX DE Pro-urokinase with signal sequence.  
 XX DE pro-urokinase (pro-UK); plasminogen activator; PUK33; ss.  
 XX FH Key Location/Qualifiers  
 XX FT Peptide 1..20  
 XX FT Protein /label=signal peptide  
 XX FT 21..431  
 XX FT /label=pro-urokinase  
 XX PN EP265874-A.  
 XX PD 04-MAY-1988.  
 XX PF 23-OCT-1987; 87EP-0115600.  
 XX PR 23-OCT-1986; 86JP-0253078.  
 XX PA (GREG ) GREEN CROSS CORP.  
 XX PI Amatsuji Y, Okabayashi K, Nagai M, Arimura H, Suyama T;  
 XX DR WPI; 1988-121000/18.  
 XX DR N-PSDB; AAN81558.  
 XX PT Glycosylated single-chain pro-urokinase prodn -  
 PT by cultivating DHFR gene-deficient CHO-K1 cells transformed  
 PT with a plasmid contg cDNA, SV40 promoter and DHFR gene.  
 XX PS Disclosure; Page ?; ?pp; English.  
 XX CC The Arg at position 2 is encoded by TGA(sic). Possible error in the  
 CC specification. Should read CGA ?  
 CC The pro-UK gene was derived from plasmid pUK33. The cDNA was  
 CC synthesised using urokinase mRNA isolated from a human kidney cell  
 CC line. Pro-UK was cloned into a SV40 promoter-contg. plasmid, down-  
 CC stream of the promoter. This plasmid was then ligated to a DHFR-  
 CC gene contg. plasmid so that pro-UK and DHFR are inserted in  
 CC opposite directions. The recombinant plasmid was used to transform  
 CC CHO-K1 cell derived DHFR gene-deficient host cells to produce  
 CC glycosylated single-chain pro-UK.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX SQ Sequence 431 AA;  
 Query Match 100.0%; Score 2301; DB 9; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-178;  
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFNIHWNCNPKFGQHCIDKSKTCYEGNGHFYRG 60  
 DB 21 SNEHQVPSNCDCLNGGTCVSNKYFNIHWNCNPKFGQHCIDKSKTCYEGNGHFYRG 80  
 QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
 DB 81 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 140  
 QY 121 PLVQECMWHDCADGKKPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPF AAIYRRH 180  
 DB 141 PLVQECMWHDCADGKKPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPF AAIYRRH 200  
 QY 181 RGSVTVYCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 240  
 DB 201 RGSVTVYCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 260  
 QY 241 LHKDYSADTLAHHNDIALKIRSKGRCAQPSRTIQTICLPMSYNDPQFSGTCEITGFGK 300

Db 261 LHKDYSADTLAHNDIALALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGK 320  
QY 301 ENSTDYLYPEQLKMTVVVKLLISHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDSSGGL 360  
Db 321 ENSTDYLYPEQLKMTVVVKLLISHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDSSGGL 380  
QY 361 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411  
Db 381 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431  
  
RESULT 11  
AAP92119  
ID AAP92119 standard; protein; 431 AA.  
XX AC AAP92119;  
XX DT 25-MAR-2003 (updated)  
XX DT 29-JUN-1990 (first entry)  
XX DE Natural human prourokinase.  
XX XX Human prourokinase; antithrombotic; derivative.  
XX XX Homo sapiens.  
XX FH Key Location/Qualifiers  
FT Misc-difference 1 /note="Optional in new deriv."  
FT Misc-difference 2..155  
FT Misc-difference 135 /note="Incorporated into new deriv."  
FT Misc-difference 156 /note="May be replaced by a non-basic AA in new deriv."  
FT Misc-difference 157 /note="Undefined residue in new deriv."  
FT Misc-difference 157 /note="Pro, Gly, Ala or Val in new deriv."  
FT Misc-difference 158 /note="Lys or Arg in new deriv."  
XX XX  
XX W08901513-A.  
XX XX  
XX 23-FEB-1989.  
XX XX  
XX 18-AUG-1988; 88WO-JP00815.  
XX XX  
XX 19-AUG-1987; 87JP-0204149.  
XX XX  
XX (SAGA ) SAGAMI CHEM RES CENTRE.  
XX (CENG ) CENTRAL GLASS CO LTD.  
XX (HODO ) HODOGAYA CHEM KK.  
XX (NIPS ) NIPPON SODA CO  
XX (NISC ) NISSAN CHEM IND LTD.  
XX XX  
XX Kobayashi Y, Omori M, Yamada C;  
XX PI  
XX WPI; 1989-068869/09.  
XX DR N-PSDB; AAN91075.  
XX XX  
XX Antithrombotic fast-acting pro-urokinase deriv. -  
XX PT produced by culture of E. coli transformant contg. new plasmid  
XX PT of PMUT9Q family.  
XX XX  
XX Disclosure; Fig 1; 75pp; Japanese.  
XX XX  
XX A human prourokinase (PU) deriv. is new which is based upon residues  
XX 2-155 of natural human prourokinase. The new deriv. is produced by  
XX E. coli J103/PMUT9Q-RPK in culture. It is a fast-acting drug for  
XX the treatment and prevention of thrombosis.  
XX (Updated on 25-MAR-2003 to correct PA field.)  
XX (Updated on 25-MAR-2003 to correct PI field.)  
XX Sequence 431 AA;

Query Match 100.0%; Score 2301; DB 10; Length 431;  
Best Local Similarity 100.0%; Pred. No. 5.7e-178;  
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHGYRG 60  
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHGYRG 80  
QY 61 KASDTDMGRPCLPWSNATVLOQTYHAHRSDALQGLGKHNYCRPNDRRRPWCYVQVGLK 120  
Db 81 KASDTDMGRPCLPWSNATVLOQTYHAHRSDALQGLGKHNYCRPNDRRRPWCYVQVGLK 140  
QY 121 PLVOECMVHDCADGKPKSSPPEELKFCQGGOKTLRPFKIIIGGEFTTIENQPFMAIYRRH 180  
Db 141 PLVOECMVHDCADGKPKSSPPEELKFCQGGOKTLRPFKIIIGGEFTTIENQPFMAIYRRH 200  
QY 181 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDIYVILGSRSLNSNTQGEKMEVENLI 240  
Db 201 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDIYVILGSRSLNSNTQGEKMEVENLI 260  
QY 241 LHKDYSADTLAHNDIALALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGK 300  
Db 261 LHKDYSADTLAHNDIALALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGK 320  
QY 301 ENSTDYLYPEQLKMTVVVKLLISHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDSSGGL 360  
Db 321 ENSTDYLYPEQLKMTVVVKLLISHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDSSGGL 380  
QY 361 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411  
Db 381 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431  
  
RESULT 12  
AAP92119  
ID AAP92119 standard; protein; 431 AA.  
XX AC AAP92119;  
XX DT 25-MAR-2003 (updated)  
XX DT 12-SEP-1990 (first entry)  
XX XX Human pro-urokinase from the cDNA of clone pcUK176.  
XX KW Non-glycosylated; pro-urokinase; E. coli; Ptip promoter; MS-2 RBS.  
XX OS Synthetic.  
XX XX  
XX EP365894-A.  
XX PD 02-MAY-1990.  
XX XX  
XX 06-OCT-1989; 89EP-0118586.  
XX XX  
XX 11-OCT-1988; 88GB-0023833.  
XX XX  
XX (FARM ) FARMITALIA ERBA SPA CARLO.  
XX XX  
XX Brandazza A, Sarmientos P, Orsini G;  
XX WPI; 1990-133447/18.  
XX DR N-PSDB; AAQ04107.  
XX XX  
XX Non-glycosylated pro-urokinase prodn. - using E.coli B strains and E.coli  
XX promoter Ptip and Shine-Dalgarno sequence MS-2.  
XX PS Disclosure; Page 2; ?pp; English.  
XX CC SER residue at position 21 is the start of the mature proUK.  
XX CC Non-glycosylated proUK (MW 45kd) produced by E.coli B strain containing  
XX the sequence..  
XX See also AAQ04101-07.



[illegible]

RESULT 15	
AA63141	
ID	AA63141 standard; Protein; 431 AA.
XX	
AC	AA63141;
XX	
DT	25-MAR-2003 (updated)
DT	09-JUN-1995 (first entry)
XX	
DE	Full length human urokinase protein.
XX	
KW	Human urokinase glycoproteins; cardiovascular diseases;
KW	pulmonary embolism.
XX	
OS	Homo sapiens.
XX	
XX	
Key	Location/Qualifiers
FT	Sig peptide
FT	1..20
FT	Disulfide-bond
FT	70..151
FT	Disulfide-bond
FT	91..133
FT	Disulfide-bond
FT	122..146
FT	Disulfide-bond
FT	168..299
FT	Disulfide-bond
FT	209..225

FT	Disulfide-bond	217..288	
FT	Disulfide-bond	313..382	
FT	Disulfide-bond	345..381	
FT	Disulfide-bond	372..400	
FT	Cleavage-site	179..180	
FT	/note= "cleavage of this site produces a bioactive two chain form of urokinase"		
XX	PN	EP620279-A1.	
XX	PD	19-OCT-1994.	
XX	PF	14-APR-1983; 94EP-0104777.	
XX	PR	15-APR-1982; 82US-0368773.	
PR	PR	14-MAR-1983; 83US-0474930.	
PR	PR	14-APR-1983; 83EP-0103629.	
XX	PA	(GETH ) GENENTECH INC.	
XX	PI	Heyneker HL, Holmes WE, Vohar GA;	
XX	DR	WPI; 1994-318362/40.	
DR	DR	N-PSDB; RAQ73483.	
XX	PT	Prodn. of human urokinase glycoproteins - using a recombinant expression system used for the treatment of vascular diseases or conditions.	
XX	PS	Claim 1; Fig 4; 4lpp; English.	
XX	CC	AAQ73483 is the cDNA sequence which encodes AAR63141 the full length 54000 dalton human urokinase (UK) protein. This cDNA was used in the construction of a plasmid capable of transforming either yeast or vertebrate cells, enabling them to produce the 54000 dalton human UK protein. The UK glycoprotein produced could then be used in the treatment of cardiovascular diseases, including pulmonary embolism. The UK produced using this method had the advantage of a specific activity towards fibrin and extant thrombi, not demonstrated previously with UK isolated from natural sources. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct PR field.)	
XX	SQ	Sequence 431 AA;	
Query Match			
Best Local Similarity 100.0%; Score 2301; DB 15; Length 431;			
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

Qy	1	SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFIYRG	60
Db	21	SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFIYRG	80
Qy	61	KASDTDTWGRPCLPNWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPCWYQVQGLK	120
Db	81	KASDTDTWGRPCLPNWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPCWYQVQGLK	140
Qy	121	PLVQECMVHDCADGCKPSSPEELKFCQGOKTLRPRFKIIGGEFTTINQPMFAAIYRRH	180
Db	141	PLVQECMVHDCADGCKPSSPEELKFCQGOKTLRPRFKIIGGEFTTINQPMFAAIYRRH	200
Qy	181	RGGSVTYVCGSLISPWCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFVEVNI	240
Db	201	RGGSVTYVCGSLISPWCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFVEVNI	260
Qy	241	LHKDYSADTTLAHHNDIALLKIRSEGRCAQPSRTIQTICLPSMYNDPQGTSCFITGPK	300
Db	261	LHKDYSADTTLAHHNDIALLKIRSEGRCAQPSRTIQTICLPSMYNDPQGTSCFITGPK	320
Qy	301	ENSTDYILYPEOLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQGSGGPL	360
Db	321	ENSTDYILYPEOLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQGSGGPL	380

QY 361 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 411  
 DB 381 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

## RESULT 16

AAV99591  
 ID AAY99591 standard; protein; 431 AA.

XX AC AAY99591;

XX DT 13-SEP-2000 (first entry)

XX DE Human plasminogen activator urokinase, u-PA.

XX KW Human; serine protease; plasminogen activator; cardiant;

XX KW Thrombolytic; heart attack; stroke; blood clotting disorder.

XX OS Homo sapiens.

XX PN WO200032759-A1.

XX PD 08-JUN-2000.

XX PF 06-MAY-1999; 99WO-US09991.

XX PR 02-DEC-1998; 98US-0110588.

XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.

XX PI Lin X, Zhang XC, Tang JUN;

XX DR WPI; 2000-422975/36.

XX PS Disclosure; Page 26-28; 41pp; English.

XX CC The present sequence is human plasminogen activator urokinase (u-PA), a serine protease which hydrolyses a peptide bond in human plasminogen to convert it to its active form, plasmin. Plasminogen is the principal serine protease zymogen in the extracellular fluids of vertebrates. Plasmin is implicated in pericellular proteolysis associated with a wide range of physiological and pathological processes. Plasminogen activators regulate plasminogen expression either by hydrolysing a peptide bond, as in the case of u-PA, or by forming tight binding complexes with plasminogen to spontaneously convert it to plasmin. Review of sequence homologies of several plasminogen activators and chymotrypsin has identified a six amino acid peptide involved in plasminogen activation. This peptide is particularly useful when inserted between amino acid residues 644 and 645 of full length human plasminogen. Novel plasminogen activators have been made based upon the plasminogen activation/recognition site of plasminogen binding proteins. The polypeptides are useful in preparing thrombolytic agents for treating blood clotting disorders such as heart attack.

XX SQ Sequence 431 AA;

Query Match 100.0%; Score 2301; DB 21; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-178;  
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHLQVPSNCDLNGGTCVSNKYFSNHWNCNPKFGGHCEDKSKTCYEGNGHYRG 60  
 DB 21 SNEHLQVPSNCDLNGGTCVSNKYFSNHWNCNPKFGGHCEDKSKTCYEGNGHYRG 80

QY 61 KASDTDMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
 DB 81 KASDTDMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140

QY 121 PLVQECMWHDCAKGKSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 180  
 DB 141 PLVQECMWHDCAKGKSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 200  
 QY 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKEFEVENLI 240  
 DB 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKEFEVENLI 260  
 QY 241 LHKDYSADTLAHHNDIALLKIRSKEGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGK 300  
 DB 261 LHKDYSADTLAHHNDIALLKIRSKEGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGK 320  
 QY 301 ENSTDYLYPEQLKMTVVKLISHRECQOPHYVGVSEVTTTMLCAADPQWKDSCQSDSGPL 360  
 DB 321 ENSTDYLYPEQLKMTVVKLISHRECQOPHYVGVSEVTTTMLCAADPQWKDSCQSDSGPL 380  
 QY 361 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 411  
 DB 381 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

## RESULT 17

AAV50869  
 ID AAY50869 standard; protein; 431 AA.

XX AC AAY50869;

XX DT 24-FEB-2000 (first entry)

XX DE Human urokinase protein fragment.

XX KW Urokinase; human; thrombolytic agent; streptokinase; antigenic;

XX KW blood clot; heart attack; treatment.

XX OS Homo sapiens.

XX PN WO9957251-A2.

XX PD 11-NOV-1999.

XX PF 06-MAY-1999; 99WO-US10086.

XX PR 06-MAY-1998; 98US-0084392.

XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.

XX PI Zhang XC, Lin X, Tang JUN;

XX DR WPI; 2000-052966/04.

XX PT New thrombolytic agents derived from modified humanized streptokinase, useful for treating blood clot disorders -

XX PS Disclosure; Page 46-48; 55pp; English.

XX CC This invention describes a novel thrombolytic agent comprising streptokinase where at least one nonessential portion has been modified. The invention also describes a method of forming a thrombolytic agent comprising determining a nonessential portion of streptokinase and modifying the nonessential portion to render the resulting protein less antigenic. The modified streptokinase is used to treat blood clot disorders, such as heart attacks. The modified streptokinase has less antigenicity than streptokinase but is still able to complex plasminogen and lead to plasminogen activation. Modified streptokinase with the nonessential portions removed or truncated simplify the molecule. Such smaller proteins are cheaper and easier to produce. This sequence represents a fragment of the human urokinase protein which is used in the description of the method of the invention.

XX SQ Sequence 431 AA;

Query Match 100.0%; Score 2301; DB 21; Length 431;





FT Peptide 1..20 /note= "Signal peptide"  
 FT Protein 21..431 /label= "Pro-urokinase/HMW-uPA"  
 FT Protein 156..431 /label= "LMW-uPA"  
 FT Protein /note= "With a cleavage between amino acids 178-179"  
 FT Protein /note= "With a cleavage between amino acids 178-179"  
 XX EPI232755-A2.  
 XX 21-AUG-2002.  
 XX 15-FEB-2002; 2002EP-0003555.  
 XX 20-FEB-2001; 2001JP-0042655.  
 XX 19-JUN-2001; 2001JP-0184284.  
 XX (JCRP-) JCR PHARM CO LTD.  
 XX Wada M, Wada N;  
 XX WPI; 2002-610512/66.  
 XX N-PSDB; ABA00207.  
 XX Anti-HIV agents, comprises ligand molecule that binds to CD87, e.g.  
 PT high molecular weight urokinase-type plasminogen activator,  
 PT amino-terminal fragment or an anti-CD87 antibody -  
 XX  
 PS Disclosure; Page 20-23; 38pp; English.  
 CC This sequence represents single chain prepro-urokinase (sc-uPA).  
 CC Pro-urokinase (amino acids 21-431) with a cleavage between amino  
 CC acids 178 and 179 gives high molecular weight urokinase-type  
 CC plasminogen activator (HMW-uPA). HMW-uPA is a protein consisting  
 CC of two peptide chains linked by a di-sulphide bond. The chains,  
 CC long A and B, are formed by enzymatic cleavage between amino acids  
 CC 178 and 179 of pro-urokinase. HMW-uPA includes an BGF-like domain,  
 CC a kringle domain and a urokinase receptor (CD87) binding domain.  
 CC HMW-uPA is then cleaved between amino acids 155 and 156 to give low  
 CC molecular weight urokinase-type plasminogen activator (LMW-uPA)  
 CC (amino acids 156-178 and 179-431), that has no plasminogen activator  
 CC activity. sc-uPA, or fragments of it, may be used in the anti-HIV  
 CC agents of the invention which comprise a ligand molecule that binds to  
 CC CD87. The agents are useful for treating HIV-infected humans for  
 CC suppression of reproduction of HIV. The anti-HIV agents act by a  
 CC mechanism of action different from those of conventional drugs, of  
 CC widening the choice of therapeutics agents and avoiding problems of  
 CC resistant HIV.  
 XX  
 SQ Sequence 431 AA;  
 Query Match 100.0%; Score 2301; DB 23; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-178;  
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNELHQVPSNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCHIDSKTCYEGNGHYRG 60  
 DB 21 SNELHQVPSNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCHIDSKTCYEGNGHYRG 80  
 QY 61 KASTDTWGRCLPWNSTVLOOTYHAHRSALQGLGKHNYCRPNDRRPWCYVQVGLK 120  
 DB 81 KASTDTWGRCLPWNSTVLOOTYHAHRSALQGLGKHNYCRPNDRRPWCYVQVGLK 140  
 QY 121 PLVCECMVHDCADGKFPSSPEELKFCGQKTLRPFKIIIGGFETTINQPFWFAALVRRH 180  
 DB 141 PLVCECMVHDCADGKFPSSPEELKFCGQKTLRPFKIIIGGFETTINQPFWFAALVRRH 200  
 QY 181 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDIIVYLGSRSLNSNTQEMKFEVENLI 240  
 DB 201 RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDIIVYLGSRSLNSNTQEMKFEVENLI 260  
 QY 241 LHKDYSADTLAHHNDIALKIRSEGRCAQPSRTIQICLPSPMYNDPQFGTSCITGFGK 300

DB 261 LHKDYSADTLAHHNDIALKIRSEGRCAQPSRTIQICLPSPMYNDPQFGTSCITGFGK 320  
 QY 301 ENSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDSDGGL 360  
 DB 321 ENSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDSDGGL 380  
 QY 361 VCSLQGRMTLTGIVSWGRGCALCKDKPGVYTRVSHFLPWIRSHRTEENGLAL 411  
 DB 381 VCSLQGRMTLTGIVSWGRGCALCKDKPGVYTRVSHFLPWIRSHRTEENGLAL 431  
 RESULT 20  
 AAU99228  
 ID AAU99228 standard; Protein; 431 AA.  
 XX  
 AC AAU99228;  
 XX  
 DT 24-SEP-2002 (first entry)  
 XX  
 XX Human plasminogen activator, urokinase (PLAU).  
 DE Human; Plasminogen activator; urokinase; PLAU; cancer; enzyme;  
 KW cytostatic; serine protease; thrombolytic disorder; isogene;  
 KW pulmonary embolism; chromosome 10q24-qter; haplotype; genotype;  
 KW SNP; single nucleotide polymorphism; thrombolytic; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200240503-A2.  
 XX  
 PD 23-MAY-2002.  
 XX  
 PF 14-NOV-2001; 2001WO-US44001.  
 XX  
 PR 17-NOV-2000; 2000US-249703P.  
 XX  
 PA (GENA-) GENAISSANCE PHARM INC.  
 XX  
 XX Anastasio AE, Bentivegna SC, Koshy B;  
 DR WPI; 2002-519370/55.  
 DR N-PSDB; ABK86597, ABK86598.  
 XX  
 PT Genetic variants of plasminogen activator, Urokinase (PLAU) isogenes,  
 PT useful for improving efficiency and reliability in drug development for  
 PT treating thrombolytic disorders and cancer -  
 XX  
 PS Claim 27; Fig 3; 92pp; English.  
 XX  
 CC The invention relates to a polynucleotide comprising a first nucleotide  
 CC sequence (NS1) comprising a PLAU (plasminogen activator, urokinase,  
 CC a serine protease) isogene selected from isogenes 1-9 and 11-20 given  
 CC in the specification, where each isogene comprises the regions of the  
 CC PLAU gene or cDNA and is further defined by the corresponding sequence of  
 CC polymorphisms (defining single nucleotide polymorphisms, SNP). Also  
 CC included are methods of haplotyping/genotyping (and predicting the  
 CC haplotype/genotype of the PLAU gene of an individual, identifying an  
 CC association between a trait and at least one haplotype or haplotype pair  
 CC of the PLAU gene, an isolated oligonucleotide for detecting a  
 CC polymorphism in the PLAU gene, a recombinant non-human organism  
 CC transformed or transfected with the gene or cDNA, fragments of the  
 CC polynucleotides of at least 10 base pairs encompassing a polymorphic  
 CC site, an isolated polymorphic variant PLAU protein or fragment, an  
 CC isolated monoclonal antibody specific for PLAU, a computer system for  
 CC storing and analysing polymorphism data for the PLAU gene and a genome  
 CC anthology for the PLAU gene. PLAU is useful in screening for drugs  
 CC targeting PLAU that are useful for treating thrombolytic disorders and  
 CC cancers. The methods are useful for improving the efficiency and  
 CC reliability of the discovery and development of drugs for treating  
 CC diseases associated with PLAU activity, in validating PLAU as a drug  
 CC target and in the design of clinical trials for treating a specific  
 CC condition of disease associated with PLAU activity. The antibody is

CC useful in diagnostic, prognostic and therapeutic methods. PLAU  
CC polynucleotides are useful in studying the expression and function of  
CC PLAU, and in expressing PLAU protein for use in screening for candidate  
CC drugs to treat diseases related to PLAU activity. The gene for PLAU  
CC is located on chromosome 10q24-qter. The present sequence represents the  
CC PLAU protein.  
XX  
XX

SQ Sequence 431 AA;

Query Match 100.0%; Score 2301; DB 23; Length 431;  
Best Local Similarity 100.0%; Pred. No. 5.7e-178;  
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60  
DB 21 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 80  
QY 61 KASTDTMGRECLPWN SATVLQQT YHAHRS DALQGLGKHNYCRPNDRRPWCYVQVGLK 120  
DB 81 KASTDTMGRECLPWN SATVLQQT YHAHRS DALQGLGKHNYCRPNDRRPWCYVQVGLK 140  
QY 121 PLVQECWVHDCADGKPKSSPPELKFQCGQKTLRPRFKIIGGEFTTIENOPWFAAIYRRH 180  
DB 141 PLVQECWVHDCADGKPKSSPPELKFQCGQKTLRPRFKIIGGEFTTIENOPWFAAIYRRH 200  
QY 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 240  
DB 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260  
QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 300  
DB 261 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320  
QY 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGSEVTTKMLCAADPQWKTDSCQDGGPL 360  
DB 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGSEVTTKMLCAADPQWKTDSCQDGGPL 380  
QY 361 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411  
DB 381 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 21

ID AAE17128  
XX AAE17128 standard; Protein; 431 AA.

AC AAE17128;

DT 18-APR-2002 (first entry)

XX Human uPA protein.

XX Human; cancer; urokinase-type plasminogen activator; uPA; inflammation;  
KW Bts-1 transcription factor; N-acetylglucosaminyltransferase V; Gnt-V;  
KW matrix-type metalloproteinase; MMP-1; MMP-3; gene therapy.

OS Homo sapiens.

XX WO200196606-A2.

XX 20-DEC-2001.

PF 14-JUN-2001; 2001WO-US19248.

XX 14-JUN-2000; 2000US-0593488.

PR (NYXI-) NYXIS NEURO THERAPIES INC.

XX Yamamoto H, Kroes R, Moskal JR;

XX WPI; 2002-130746/17.

DR N-PSDB; AAD27855.

XX

PT

PT transcription factor Bts-1, N-acetylglucosaminyltransferase V,  
PT urokinase-type plasminogen activator, matrix-type metalloproteinase-1  
PT and -3 gene expression -  
XX

PS Example 1; Page 62-63; 63pp; English.

XX The invention relates to a method of identifying a compound for treating  
CC cancer. The method involves detecting the expression of a panel of  
CC sequences selected from transcription factor Bts-1, urokinase-type  
CC plasminogen activator (uPA), N-acetylglucosaminyltransferase V (Gnt-V),  
CC matrix-type metalloproteinase (MMP)-1 and MMP-3 in the cell. The method  
CC is useful for identifying a compound that affects a cell, particularly a  
CC cancer cell or glioma cell, or a cell that is involved in inflammation.  
CC It is used for diagnosing and/or treating cancer or other conditions that  
CC are affected by one or more members of a panel of genes or their protein  
CC product. The method is also useful for drug discovery, drug safety  
CC evaluations and in gene therapy. The present sequence is human uPA  
CC protein.  
XX

SQ Sequence 431 AA;

Query Match 100.0%; Score 2301; DB 23; Length 431;  
Best Local Similarity 100.0%; Pred. No. 5.7e-178;  
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60  
DB 21 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 80  
QY 61 KASTDTMGRECLPWN SATVLQQT YHAHRS DALQGLGKHNYCRPNDRRPWCYVQVGLK 120  
DB 81 KASTDTMGRECLPWN SATVLQQT YHAHRS DALQGLGKHNYCRPNDRRPWCYVQVGLK 140  
QY 121 PLVQECWVHDCADGKPKSSPPELKFQCGQKTLRPRFKIIGGEFTTIENOPWFAAIYRRH 180  
DB 141 PLVQECWVHDCADGKPKSSPPELKFQCGQKTLRPRFKIIGGEFTTIENOPWFAAIYRRH 200  
QY 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 240  
DB 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260  
QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 300  
DB 261 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320  
QY 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGSEVTTKMLCAADPQWKTDSCQDGGPL 360  
DB 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGSEVTTKMLCAADPQWKTDSCQDGGPL 380  
QY 361 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411  
DB 381 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 22

ID ABUS6547  
XX ABUS6547 standard; Protein; 431 AA.

AC ABUS6547;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polypeptide #140.

XX Lung cancer-associated polypeptide; cytostatic; emphysema;  
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

OS Unidentified.

XX



XX Sequence 431 AA;  
XX Query Match 100.0%; Score 2301; DB 24; Length 431;  
XX Best Local Similarity 100.0%; Pred. No. 5.7e-178;  
XX Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
DB 21 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80  
QY 61 KASDTMTGRPCLPWNSATVLIQOTYVHAHRSALQGLGKHNYCRPNDRRRPWCYVQVGLK 120  
DB 81 KASDTMTGRPCLPWNSATVLIQOTYVHAHRSALQGLGKHNYCRPNDRRRPWCYVQVGLK 140  
QY 121 PLVQECMVHDCADGKPKSSPPEELKFQCGQKTLRPRFKIIGGEPTTIENTOPWFAAIYRRH 180  
DB 141 PLVQECMVHDCADGKPKSSPPEELKFQCGQKTLRPRFKIIGGEPTTIENTOPWFAAIYRRH 200  
QY 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKEDYIVVLSRSLNSNTQGMKEFEVENLI 240  
DB 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKEDYIVVLSRSLNSNTQGMKEFEVENLI 260  
QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCAPSRITQITCLPSMYNDPQFGTSCBITGFGK 300  
DB 261 LHKDYSADTLAHHNDIALLKIRSKGRCAPSRITQITCLPSMYNDPQFGTSCBITGFGK 320  
QY 301 ENSTDYLYPEQLKMTVVVKLISHRECOQPHYVYGSEVTTMLCAADPQWKTDSCQDSSGGL 360  
DB 321 ENSTDYLYPEQLKMTVVVKLISHRECOQPHYVYGSEVTTMLCAADPQWKTDSCQDSSGGL 380  
QY 361 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKENGLAL 411  
DB 381 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKENGLAL 431  
RESULT 24  
ABU11076  
ID ABU11076 standard; Protein; 431 AA.  
XX AC ABU11076;  
XX DT 05-FEB-2003 (first entry)  
XX DE Human urokinase plasminogen activator.  
XX KW Urokinase plasminogen activator; gene therapy; cancer;  
XX KW hyperproliferative disorder; cancer; breast cancer; colon cancer;  
XX KW bone cancer; brain cancer; ovary cancer; cervix cancer;  
XX KW endometrium cancer; stomach cancer; kidney cancer; tumour metastasis.  
XX OS Homo sapiens.  
XX PN WO200279515-A1.  
XX PD 10-OCT-2002.  
XX PF 18-MAR-2002; 2002WO-US08112.  
XX PR 30-MAR-2001; 2001US-0821972.  
XX PA (ISIS-) ISIS PHARM INC.  
XX PI Baker BF, Freier SM, Watt AT;  
XX DR WPI: 2003-058441/05.  
XX DR N-PSDB; ABX17681.  
XX PT New antisense compound, useful for preparing a composition for treating  
PT hyperproliferative disorders, cancer e.g., breast, colon, bone, brain,  
PT ovary, cervix, endometrium, stomach or kidney cancer, or tumor  
PT metastasis -  
XX

PS Disclosure; Page 101-102; 153pp; English.  
XX A new compound, which is 8-50 nucleobases in length targeted  
CC to a nucleic acid molecule encoding urokinase plasminogen activator,  
CC specifically hybridises with and inhibits the expression of urokinase  
CC plasminogen activator. The compound is useful for preparing a  
CC composition for treating (e.g. by gene therapy) hyperproliferative  
CC disorder, cancer e.g., breast, colon, bone, brain, ovary, cervix,  
CC endometrium, stomach or kidney cancer, or tumour metastasis. This  
CC is the amino acid sequence of a urokinase plasminogen activator.  
XX Sequence 431 AA;  
XX Query Match 100.0%; Score 2301; DB 24; Length 431;  
XX Best Local Similarity 100.0%; Pred. No. 5.7e-178;  
XX Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
DB 21 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80  
QY 61 KASDTMTGRPCLPWNSATVLIQOTYVHAHRSALQGLGKHNYCRPNDRRRPWCYVQVGLK 120  
DB 81 KASDTMTGRPCLPWNSATVLIQOTYVHAHRSALQGLGKHNYCRPNDRRRPWCYVQVGLK 140  
QY 121 PLVQECMVHDCADGKPKSSPPEELKFQCGQKTLRPRFKIIGGEPTTIENTOPWFAAIYRRH 180  
DB 141 PLVQECMVHDCADGKPKSSPPEELKFQCGQKTLRPRFKIIGGEPTTIENTOPWFAAIYRRH 200  
QY 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKEDYIVVLSRSLNSNTQGMKEFEVENLI 240  
DB 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKEDYIVVLSRSLNSNTQGMKEFEVENLI 260  
QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCAPSRITQITCLPSMYNDPQFGTSCBITGFGK 300  
DB 261 LHKDYSADTLAHHNDIALLKIRSKGRCAPSRITQITCLPSMYNDPQFGTSCBITGFGK 320  
QY 301 ENSTDYLYPEQLKMTVVVKLISHRECOQPHYVYGSEVTTMLCAADPQWKTDSCQDSSGGL 360  
DB 321 ENSTDYLYPEQLKMTVVVKLISHRECOQPHYVYGSEVTTMLCAADPQWKTDSCQDSSGGL 380  
QY 361 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKENGLAL 411  
DB 381 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKENGLAL 431  
RESULT 25  
AAR20537  
ID AAR20537 standard; Protein; 434 AA.  
XX AC AAR20537;  
XX DT 25-MAR-2003 (updated)  
XX DT 21-MAY-1992 (first entry)  
XX DE Amidated deriv. of pro-urokinase (2).  
XX KW Pro-urokinase; plasminogen activator; ss.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX FT Protein 1..431  
XX FT /label= pro-UK  
XX DE4122688-A.  
XX PN 16-JAN-1992.  
XX PD 09-JUL-1991; 91DE-4122688.  
XX PR 12-JUL-1990; 90GB-0015369.  
XX PR 10-JUL-1991; 91GB-0014846.

XX (FARM ) FARMITALIA ERBA SRL CARLO.  
 XX Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;  
 XX WPI; 1992-025815/04.  
 XX New amidated derivs. of human pro-urokinase - are fibrinolytic  
 PT and can be used to treat acute myocardial infarction, pulmonary  
 PT embolism or deep venous thrombosis  
 XX Claim 4,8; Page 8; 18pp; German.  
 XX The protein has fibrinolytic activity and can be used in the same  
 CC way as PUK, e.g. for treating acute myocardial infarction, lung  
 CC embolism and deep venous thrombosis. It has greater affinity for  
 CC plasminogen bound to fibrin than for circulating plasminogen, so  
 CC have high selectivity for thrombi with reduced chance of bleeding.  
 CC Compared with the COOH-terminated cpds., it has a better stability  
 CC against most carboxypeptidases and prolonged half life.  
 CC See also AAQ20360, AAQ20754-58 and AAR20536-38.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX Sequence 434 AA;  
 SQ Query Match 100.0%; Score 2301; DB 13; Length 434;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-178;  
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
 Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80  
 QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
 Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140  
 QY 121 PLVQECWHDCAADGKKPSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFWFAIYRRH 180  
 Db 141 PLVQECWHDCAADGKKPSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFWFAIYRRH 200  
 QY 181 RGSVTVYCGGSLISPCWVISATHCFIDYPKEDYIVLGRSRLNSNTQGMKFEVENLI 240  
 Db 201 RGSVTVYCGGSLISPCWVISATHCFIDYPKEDYIVLGRSRLNSNTQGMKFEVENLI 260  
 QY 241 LHKDYSADTLAHHNDIALLKIRSEKGRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGK 300  
 Db 261 LHKDYSADTLAHHNDIALLKIRSEKGRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGK 320  
 QY 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGEVTTKMLCAADPQWKTDSCQDSDGGL 360  
 Db 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGEVTTKMLCAADPQWKTDSCQDSDGGL 380  
 QY 361 VCSLQGRMTLTGIVSWGRGCALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411  
 Db 381 VCSLQGRMTLTGIVSWGRGCALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431  
 RESULT 26  
 AAR20538  
 ID AAR20538 standard; Protein; 434 AA.  
 XX AAR20538;  
 XX AC  
 XX 25-MAR-2003 (updated)  
 DT 21-MAY-1992 (first entry)  
 XX Amidated deriv. of pro-urokinase (3).  
 DE Pro-urokinase; plasminogen activator; ss.  
 XX Homo sapiens.  
 OS

Key Location/Qualifiers  
 Protein 1..431  
 /label= pro-UK  
 DB4122688-A.  
 16-JAN-1992.  
 09-JUL-1991; 91DE-4122688.  
 12-JUL-1990; 90GB-0015369.  
 10-JUL-1991; 91GB-0014846.  
 (FARM ) FARMITALIA ERBA SRL CARLO.  
 Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;  
 WPI; 1992-025815/04.  
 New amidated derivs. of human pro-urokinase - are fibrinolytic  
 and can be used to treat acute myocardial infarction, pulmonary  
 embolism or deep venous thrombosis  
 Claim 4,9; Page 8; 18pp; German.  
 The protein has fibrinolytic activity and can be used in the same  
 way as PUK, e.g. for treating acute myocardial infarction, lung  
 embolism and deep venous thrombosis. It has greater affinity for  
 plasminogen bound to fibrin than for circulating plasminogen, so  
 have high selectivity for thrombi with reduced chance of bleeding.  
 Compared with the COOH-terminated cpds., it has a better stability  
 against most carboxypeptidases and prolonged half life.  
 See also AAQ20360, AAQ20754-58 and AAR20536-38.  
 (Updated on 25-MAR-2003 to correct PA field.)  
 Sequence 434 AA;  
 Query Match 100.0%; Score 2301; DB 13; Length 434;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-178;  
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
 Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80  
 QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
 Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140  
 QY 121 PLVQECWHDCAADGKKPSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFWFAIYRRH 180  
 Db 141 PLVQECWHDCAADGKKPSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFWFAIYRRH 200  
 QY 181 RGSVTVYCGGSLISPCWVISATHCFIDYPKEDYIVLGRSRLNSNTQGMKFEVENLI 240  
 Db 201 RGSVTVYCGGSLISPCWVISATHCFIDYPKEDYIVLGRSRLNSNTQGMKFEVENLI 260  
 QY 241 LHKDYSADTLAHHNDIALLKIRSEKGRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGK 300  
 Db 261 LHKDYSADTLAHHNDIALLKIRSEKGRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGK 320  
 QY 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGEVTTKMLCAADPQWKTDSCQDSDGGL 360  
 Db 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGEVTTKMLCAADPQWKTDSCQDSDGGL 380  
 QY 361 VCSLQGRMTLTGIVSWGRGCALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411  
 Db 381 VCSLQGRMTLTGIVSWGRGCALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431  
 RESULT 27  
 AAR20536  
 ID AAR20536 standard; Protein; 436 AA.

XX AC AAR20536;  
XX 25-MAR-2003 (updated)  
XX 21-MAY-1992 (first entry)  
XX Amidated deriv. of pro-urokinase (1).  
XX Pro-urokinase; plasminogen activator; ss.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
XX Protein 1..431  
XX /label= pro-UK  
XX /note= "the amidated form of pro-UK is  
XX claimed in claim 1"  
XX Misc-difference 433..436  
XX /note= "may be any amino acid, pref. Lys or Arg,  
XX or 0-4 amino acids may be omitted"  
XX DE4122688-A.  
XX 16-JAN-1992.  
XX 09-JUL-1991; 91DE-4122688.  
XX 12-JUL-1990; 90GB-0015369.  
XX 10-JUL-1991; 91GB-0014846.  
XX (FARM ) FARMITALIA ERBA SRL CARLO.  
XX Gozzini L, Visco C, Perego R, Roncucci R, Sammientos P;  
XX WPI; 1992-025815/04.  
XX New amidated derivs. of human pro-urokinase - are fibrinolytic  
XX and can be used to treat acute myocardial infarction, pulmonary  
XX embolism or deep venous thrombosis  
XX Claim 1,4,7; Page 8; 18pp; German.  
XX The protein has fibrinolytic activity and can be used in the same  
XX way as PUK, e.g. for treating acute myocardial infarction, lung  
XX embolism and deep venous thrombosis. It has greater affinity for  
XX plasminogen bound to fibrin than for circulating plasminogen, so  
XX have high selectivity for thrombi with reduced chance of bleeding.  
XX Compared with the COOH-terminated cpds., it has a better stability  
XX against most carboxypeptidases and prolonged half life.  
XX See also AAO20360, AAO20754-58 and AAR20536-38.  
XX (Updated on 25-MAR-2003 to correct PA field.)  
XX Sequence 436 AA;  
Query Match 100.0%; Score 2301; DB 13; Length 436;  
Best Local Similarity 100.0%; Pred. No. 5.8e-178;  
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHVPNSCCLNGTCTVSNKYFSNIHWCNPKKFGGHCIDKSKTCYEGNGHYRG 60  
DB 21 SNELHVPNSCCLNGTCTVSNKYFSNIHWCNPKKFGGHCIDKSKTCYEGNGHYRG 80  
QY 61 KASTDTWGRPCLPWNASATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRRPMCVYQVGLK 120  
DB 81 KASTDTWGRPCLPWNASATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRRPMCVYQVGLK 140  
QY 121 PLVQECWHDADCKGKPSPPPELKFQCCQKTLRPFKLIIGGEFTTIENQPFWFAIYRRH 180  
DB 141 PLVQECWHDADCKGKPSPPPELKFQCCQKTLRPFKLIIGGEFTTIENQPFWFAIYRRH 200  
QY 181 RGSSTVYVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENLI 240  
DB 201 RGSSTVYVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENLI 260

QY 241 LHKDYSADTLAHHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPQGTSCITGFGK 300  
DB 261 LHKDYSADTLAHHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPQGTSCITGFGK 320  
QY 301 ENSTDYLYPEQLKMTVVKLISHRECQPHYGYSEVTTKMLCAADPQWKTDSGQSGGGL 360  
DB 321 ENSTDYLYPEQLKMTVVKLISHRECQPHYGYSEVTTKMLCAADPQWKTDSGQSGGGL 380  
QY 361 VCSLOGRMTLTGIVSGRGCAKDKPGVYTVRVSHPLPWIRSHRTKEENGIAL 411  
DB 381 VCSLOGRMTLTGIVSGRGCAKDKPGVYTVRVSHPLPWIRSHRTKEENGIAL 431  
RESULT 28  
AAW24578  
ID AAW24578 standard; Protein; 430 AA.  
XX AAW24578;  
XX 25-MAR-2003 (updated)  
XX 11-NOV-1997 (first entry)  
XX Inhibitor resistant urokinase.  
XX Urokinase; inhibitor resistant; plasminogen; human; whey acid protein;  
XX plasminogen activator inhibitor-1; plasminogen activator; protease; WAP;  
XX plasmin; proteolytic enzyme; clot lysis; heart attack; heart muscle.  
XX Homo sapiens.  
XX Synthetic.  
XX Key Location/Qualifiers  
XX Peptide 1..19  
XX /note= "whey acid protein signal peptide"  
XX Protein 20..430  
XX /note= "urokinase"  
XX Domain 20..64  
XX /note= "B-domain"  
XX Domain 61..150  
XX /note= "Kringle-1 domain"  
XX Domain 179..430  
XX /note= "P-domain"  
XX Misc-difference 198..203  
XX /note= "deleted in modified urokinase of the invention"  
US5648253-A.  
15-JUL-1997.  
XX 08-SEP-1992; 92US-0942157.  
XX 20-DEC-1990; 90US-0631673.  
XX 08-SEP-1992; 92US-0942157.  
XX (TSIT-) TSI CORP.  
XX Wei C;  
XX WPI; 1997-372062/34.  
XX N-PSDB; AAT80075.  
XX Deletion-modified urokinase protein - with increased resistance to  
XX inhibition by plasminogen activator inhibitor-1  
XX Disclosure; Column 15-18; 16pp; English.  
XX This sequence represents the full length urokinase, including the  
XX whey acid protein (WAP) signal peptide. This sequence has residues  
XX 179-184 of the urokinase sequence deleted to create the modified  
XX urokinase of the invention. The modified urokinase (see AAW24579)  
XX cleaves plasminogen, and has a lower binding affinity for plasminogen  
XX activator inhibitor-1 than the corresponding unmodified urokinase.

CC Urokinase is one of two types of mammalian plasminogen activators (PA),  
 CC the other being tissue type PA. PAs catalyse the conversion of the  
 CC circulating zymogen plasminogen to the broad spectrum protease plasmin by  
 CC limited proteolysis. The modified urokinase can be used for clot lysis,  
 CC specifically to dissolve heart attack-causing clots before they cause  
 CC permanent damage to heart muscle. The urokinase mutant is more resistant  
 CC to inhibition by plasminogen activator inhibitors than the unmodified  
 CC urokinase. It can be selectively expressed and secreted from the mammary  
 CC glands of transgenic animals.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 XX Sequence 430 AA;

Query Match 99.9%; Score 2298; DB 18; Length 430;  
 Best Local Similarity 99.8%; Pred. No. 9.9e-178;  
 Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60  
 DB 20 SNELHQPNSDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 79  
 QY 61 KASTDTMRGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
 DB 80 KASTDTMRGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 139  
 QY 121 PLVQECMVHDCADGKKPSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 180  
 DB 140 PLVQECMVHDCADGKKPSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 199  
 QY 181 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSQTGEMKFEVENLI 240  
 DB 200 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSQTGEMKFEVENLI 259  
 QY 241 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCITGFGK 300  
 DB 260 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCITGFGK 319  
 QY 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDGGGGL 360  
 DB 320 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDGGGGL 379  
 QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGIAL 411  
 DB 380 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGIAL 430

RESULT 29  
 AAP91886  
 ID AAP91886 standard; protein; 431 AA.  
 XX  
 AC AAP91886;  
 XX

DT 25-MAR-2003 (updated)  
 DT 31-OCT-2002 (updated)  
 DT 16-APR-1990 (first entry)  
 XX

DE Sequence of prourokinase.

XX Low mol. wt. plasminogen activator; prourokinase; dissolve blood clot.  
 XX Homo sapiens.

XX Key Location/Qualifiers  
 FT Protein 21..431  
 FT /note= "Mature prourokinase."  
 FT Region 170..179  
 FT /note= "Preferred initiation region for the low mol. wt.  
 FT plasminogen activators."  
 XX

PN EP316068-A.  
 XX  
 PD 17-MAY-1989.  
 XX

PF 07-OCT-1988; 88EP-0309417.

PR 09-OCT-1987; 87US-0107370.

PR 27-SEP-1988; 88US-0248727.

XX (COLB ) COLLABORATIVE RES INC.

XX Mao JI;

XX WPI; 1989-146601/20.

DR N-PSDB; AAN91740.

XX Modified low mol. wt. plasminogen activator- formed of amino acids  
 comprising the amino acid portion of prourokinase from 150 to 411

XX Fig 1; Page -; 27pp; English.

XX A low mol. wt. polypeptide plasminogen activator (PA) is claimed which is  
 formed from amino acids 150-411 of prourokinase. The preferred initiation  
 region for the low mol. wt. PA is indicated (see Ft). A low mol wt. PA  
 can be injected into blood in the body in vivo to dissolve clots  
 without harm.

CC (Updated on 31-OCT-2002 to add missing OS field.)

CC (Updated on 25-MAR-2003 to correct PR field.)

CC (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 431 AA;

Query Match 99.9%; Score 2298; DB 10; Length 431;  
 Best Local Similarity 99.8%; Pred. No. 9.9e-178;  
 Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60

DB 21 SNELHQPNSDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMRGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120

DB 81 KASTDTMRGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKKPSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 180

DB 141 PLVQECMVHDCADGKKPSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 200

QY 181 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSQTGEMKFEVENLI 240

DB 201 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSQTGEMKFEVENLI 260

QY 241 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCITGFGK 300

DB 261 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCITGFGK 320

QY 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDGGGGL 360

DB 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDGGGGL 380

QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGIAL 411

DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGIAL 431

RESULT 30  
 AAP94764  
 ID AAP94764 standard; protein; 431 AA.  
 XX  
 AC AAP94764;  
 XX

DT 25-MAR-2003 (updated)  
 DT 27-JUN-1990 (first entry)  
 XX

DE Non-glycosylated prourokinase.

XX Prourokinase; CGE 195; plasminogen activator; blood clot lysis;

Search completed: December 3, 2003, 14:39:12  
Job time : 67.7422 secs

```
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Region 1..20
XX FT /label= signal sequence
XX
XX EP299706-A.
XX
XX PD 18-JAN-1989.
XX
XX PF 12-JUL-1988; 88EP-0306334.
XX
XX PR 13-JUL-1987; 87US-0072426.
XX PR 29-JUN-1988; 88US-0211279.
XX
XX PA (COLB ) COLLABORATIVE RES INC.
XX
XX PI Baltimore D, Moir DT, Broeze RJ;
XX
XX DR WPI; 1989-017204/03.
XX DR N-PSDB; AAN93079.
XX
XX PT New non-glycosylated, secreted plasminogen activator - pref. with
XX PT asparagine replaced or deleted, useful for treating blood clots,
XX PT expressed in non-mammalian cells.
XX
XX PS Disclosure; Page -: pp; English.
XX
XX CC myocardial infarction.
XX CC DNA encoding the protein was sequenced from plasmid pCGE195, a subclone
XX CC of two inserts isolated by screening a cDNA library prepd. from kidney
XX CC cell RNA. One of the original inserts, clone CGP31 (tag c) started in
XX CC the middle of the signal sequence. Mutants of the sequence, pref. in
XX CC which gcc (Ala) replaces aat (Asn) at nucleotides 1002-1004 (residue 302)
XX CC are used to transform hosts for the prodn. of non-glycosylated
XX CC prourokinase.
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX SQ Sequence 431 AA;
XX
Query Match 99.9%; Score 2298; DB 10; Length 431;
Best Local Similarity 99.8%; Pred. No. 9.9e-178;
Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHVPNSCCLNGGTCVSNKIFSNHWCNCPKFKGGQHCEIDKSKTCYEGNGHFYRG 60
Db 21 SNELHVPNSCCLNGGTCVSNKIFSNHWCNCPKFKGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLFWSATVLQQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVOVGLK 120
Db 81 KASTDTMGRPCLFWSATVLQQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVOVGLK 140
QY 121 PLVQECWHDGADGKPPSPPELKFQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRH 180
Db 141 PLVQECWHDGADGKPPSPPELKFQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRH 200
QY 181 RGSVTVYVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFVEVNI 240
Db 201 RGSVTVYVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFVEVNI 260
QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCQAPSRITQITCLPSMYNDPQFGTSCEITGFGK 300
Db 261 LHKDYSADTLAHHNDIALLKIRSKGRCQAPSRITQITCLPSMYNDPQFGTSCEITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYGSEVTTKMLCAADPQWKTDSCQGDGGPL 360
Db 321 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYGSEVTTKMLCAADPQWKTDSCQGDGGPL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFLPWIRSHTKENGIAL 411
Db 381 LCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFLPWIRSHTKENGIAL 431
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:58 ; Search time 21.2586 Seconds  
(without alignments)  
1859.261 Million cell updates/sec

Title: US-09-880-503-3  
Perfect score: 2301  
Sequence: 1 SNELHQVPSNCDLNGTGV.....VSHFLPWIRSHTRKENGIAL 411

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2298	99.9	431	1 UKHU	u-plasminogen acti
2	2148	93.4	433	1 UKBAY	u-plasminogen acti
3	1859.5	80.8	442	1 UKPG	u-plasminogen acti
4	1767	76.8	433	1 JN0560	u-plasminogen acti
5	1690.5	73.5	432	1 S18932	u-plasminogen acti
6	1660.5	72.2	433	1 UKMS	u-plasminogen acti
7	1004	43.6	434	1 A35005	u-plasminogen acti
8	864.5	37.6	477	2 JS0598	t-plasminogen acti
9	864.5	37.6	562	1 UKHUT	t-plasminogen acti
10	863.5	37.5	477	1 A34369	t-plasminogen acti
11	858.5	37.3	431	2 JS0599	t-plasminogen acti
12	844.5	36.7	559	1 A35029	t-plasminogen acti
13	836.5	36.4	477	2 JS0597	t-plasminogen acti
14	831.5	36.1	559	1 A29941	t-plasminogen acti
15	752	32.7	334	2 JS0600	t-plasminogen acti
16	735.5	32.0	655	1 A46888	hepatocyte growth
17	718.5	31.2	603	2 S28941	coagulation factor
18	692	30.1	615	1 KFHU12	coagulation factor
19	674.5	29.3	558	2 JC5878	plasma hyaluronan-
20	661.5	28.7	560	1 JC4795	plasma hyaluronan-
21	642	27.9	593	2 S45281	coagulation factor
22	507.5	22.1	460	2 B16345	plasmin (EC 3.4.21
23	504.5	21.9	790	1 PLPG	plasmin (EC 3.4.21
24	501	21.8	810	1 PLHU	plasmin (EC 3.4.21
25	497	21.6	810	2 B30848	plasmin (EC 3.4.21
26	497	21.6	812	1 PLMS	plasmin (EC 3.4.21
27	484.5	21.1	812	1 PLBO	plasmin (EC 3.4.21
28	484	21.0	1420	2 A32869	apolipoprotein(a)
29	474	20.6	4548	1 S00657	apoprotein(a) (EC

30	457.5	19.9	455	2	A61545	plasmin (EC 3.4.21
31	445	19.3	761	2	JC5759	brain-specific ser
32	435	18.9	810	2	I46260	plasmin (EC 3.4.21
33	410	17.8	638	1	QKMSPL	plasma kallikrein
34	408.5	17.8	343	1	A57014	proctasin (EC 3.4.
35	406	17.6	711	1	A47136	macrophage-stimula
36	402.5	17.5	417	1	S00845	hepsin (EC 3.4.21.
37	402	17.5	638	1	KQHUP	plasma kallikrein
38	401	17.4	416	1	J33777	hepsin (EC 3.4.21.
39	401	17.4	855	2	JC7731	membrane-bound arg
40	395.5	17.2	248	2	S55066	trypsin (EC 3.4.21
41	394.5	17.1	229	1	TRBOTR	trypsin (EC 3.4.21
42	394.5	17.1	716	1	JC5061	macrophage-stimula
43	391.5	17.0	716	1	A40352	macrophage-stimula
44	390.5	17.0	247	2	S13813	trypsin (EC 3.4.21
45	390	16.9	263	2	A21195	chymotrypsin (EC 3

ALIGNMENTS

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human  
N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminog  
N;Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen a  
in form  
C;Species: Homo sapiens (man)  
C;Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000  
C;Accession: A00931; I52209; J70102; A37561; I38102; S65783; A37562; A37563; A37564; A  
R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastiao, G.; Boast, S.; Blasi, F.  
Nucleic Acids Res. 13, 2759-2771, 1985  
A;Title: The human urokinase-plasminogen activator gene and its promoter.  
A;Reference number: A00931; MUID:85215647; PMID:2987867  
A;Accession: A00931  
A;Molecule type: DNA  
A;Residues: 1-431 <RTC>  
A;Cross-references: GB:X02419; NID:937601; PIDN:CAA26268.1; PID:g1834524  
A;Note: the authors translated the codon ATG for residue 214 as Ile  
R;Nagamine, Y.; Pearson, D.; Grattan, M.  
Biochem. Biophys. Res. Commun. 132, 563-569, 1985  
A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porc  
A;Reference number: I52209; MUID:86050639; PMID:3933505  
A;Accession: I52209  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 145-161 <NAG1>  
A;Cross-references: GB:X03027; NID:9340174; PIDN:AAA61257.1; PID:g340175  
R;Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama  
Gene 36, 183-188, 1985  
A;Title: Molecular cloning of cDNA coding for human preprourokinase.  
A;Reference number: J70102; MUID:86056954; PMID:2415429  
A;Accession: J70102  
A;Molecule type: mRNA  
A;Residues: 1-213, 'I', 215-431 <NAG2>  
A;Cross-references: GB:X03226; NID:9340155; PIDN:AA097138.1; PID:g340158; GB:D00244; N  
R;Verde, P.; Stoppelli, M.F.; Galeffi, P.; Di Nocera, P.; Blasi, F.  
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984  
A;Title: Identification and primary sequence of an unspliced human urokinase poly(A) +  
A;Reference number: A37561; MUID:84272706; PMID:6589620  
A;Accession: A37561  
A;Molecule type: mRNA  
A;Residues: 66-431 <VER>  
A;Cross-references: GB:D00244; NID:9220138  
R;Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Else  
DNA 4, 139-146, 1985  
A;Title: Molecular cloning, sequencing, and expression in Escherichia coli of human pr  
A;Reference number: I38102; MUID:85203359; PMID:3888571  
A;Accession: I38102  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>  
A;Cross-references: EMBL:X02760; NID:935297; PIDN:CAA26535.1; PID:g35298

R;Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996  
 A;Title: Characterization of single chain urokinase-type plasminogen activator with a no  
 A;Reference number: S65783; MUID:96186279; PMID:8652631  
 A;Accession: S65783  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>  
 A;Cross-references: EMBL:D11143; NID:gl311467; PIDN:BARA01919.1; PID:gl199928  
 R;Gunzler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L.  
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982  
 A;Title: The primary structure of high molecular mass urokinase from human urine.  
 A;Reference number: A37562; MUID:83055084; PMID:6754569  
 A;Accession: A37562  
 A;Molecule type: protein  
 A;Residues: 21-177 <GUN>  
 R;Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O.  
 Eur. J. Biochem. 125, 251-257, 1982  
 A;Title: Human low-molecular-weight urinary urokinase. Partial characterization and prel  
 A;Reference number: A37563; MUID:83003608; PMID:6749491  
 A;Accession: A37563  
 A;Molecule type: protein  
 A;Residues: 156-176; 179-193, 'T', 195, 'T', 197-224 <SCH>  
 R;Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Flohe, L.  
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982  
 A;Title: The complete amino acid sequence of low molecular mass urokinase from human uri  
 A;Reference number: A37564; MUID:83055099; PMID:6754572  
 A;Accession: A37564  
 A;Molecule type: protein  
 A;Residues: 158-410 <STE>  
 R;Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K.  
 Biochem. Biophys. Res. Commun. 171, 401-406, 1990  
 A;Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinan  
 A;Reference number: A35689; MUID:90365737; PMID:2393398  
 A;Accession: A35689  
 A;Molecule type: protein  
 A;Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>  
 A;Note: identification of a fucose and attempt to determine its attachment site  
 R;Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzm  
 Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990  
 A;Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li  
 A;Reference number: A36697; MUID:91097529; PMID:2125213  
 A;Accession: A36697  
 A;Molecule type: protein  
 A;Residues: 21-34 <RAD>  
 R;Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.  
 submitted to the Brookhaven Protein Data Bank, July 1993  
 A;Reference number: A51255; PDB:1KDU  
 A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
 R;Li, X.; Smith, R.A.G.; Dobson, C.M.  
 Biochemistry 31, 9562-9571, 1992  
 A;Title: Sequential (1)H-NMR assignments and secondary structure of the kringle domain f  
 A;Reference number: A44375; MUID:93003110; PMID:1327118  
 A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR  
 R;Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazar, A.P.; Olejniczak,  
 submitted to the Brookhaven Protein Data Bank, January 1994  
 A;Reference number: A66822; PDB:1URK  
 A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
 R;Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;  
 submitted to the Brookhaven Protein Data Bank, July 1995  
 A;Reference number: A66058; PDB:1LMW  
 A;Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426  
 C;Comment: This enzyme is found in urine in a high molecular mass form, consisting of A  
 C;Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a  
 C;Genetics:  
 A;Gene: GDB:PTAU  
 A;Cross-references: GDB:119497; OMIM:191840  
 A;Map position: 10q24-10q24  
 A;Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3  
 C;Function:  
 A;Description: proteolytically activates plasminogen  
 A;Pathway: fibrinolysis  
 C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C;Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protei  
 F;1-20/Domain: signal sequence #status predicted <SIG>  
 F;21-431/Product: urokinase-type plasminogen activator, single chain form #status pred  
 F;21-177/Product: urokinase-type plasminogen activator chain A #status experimental <M  
 F;31-62/Domain: EGF homology <EGF>  
 F;70-151/Domain: kringle homology <KRG>  
 F;156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental  
 F;179-431/Product: urokinase-type plasminogen activator chain B #status experimental <  
 F;179-419/Domain: trypsin homology <TRY>  
 F;31-39, 33-51, 53-62, 70-151, 91-133, 122-146, 168-299, 209-225, 217-288, 313-382, 345-361, 372-  
 F;38/Binding site: carbohydrate (Thr) (covalent) #status predicted  
 F;178-179/Cleavage site: Lys-Ile (plasmin) #status experimental  
 F;224, 275, 376/Active site: His, Asp, Ser #status experimental  
 F;322/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match	99.9%	Score 2298;	DB 1;	Length 431;
Best Local Similarity	99.8%	Pred. No. 9.6e-170;		
Matches 410;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFFYRG	60
DB	21	SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFFYRG	80
QY	61	KASDTWGRPCLPWNSATVLOQTYHAHRSALQGLGHNYCRPNDRRRPWCYVQVGLK	120
DB	81	KASDTWGRPCLPWNSATVLOQTYHAHRSALQGLGHNYCRPNDRRRPWCYVQVGLK	140
QY	121	PLVOECMVHDCADGKPKSSPPEELKFCOGQKTLAPRFKIIGGEFTTIENQWFAAIYRRH	180
DB	141	PLVOECMVHDCADGKPKSSPPEELKFCOGQKTLAPRFKIIGGEFTTIENQWFAAIYRRH	200
QY	181	RGGSVTYVCGSLSPCWVISATHCIFIDYPKKEDIYVILGRSLNSNTQGEKKEVENLI	240
DB	201	RGGSVTYVCGSLSPCWVISATHCIFIDYPKKEDIYVILGRSLNSNTQGEKKEVENLI	260
QY	241	LHKDYSADTLAHHNDIALLKIRSEKRCAPSRITOTICLPSMNDPOFGTSCITGFGK	300
DB	261	LHKDYSADTLAHHNDIALLKIRSEKRCAPSRITOTICLPSMNDPOFGTSCITGFGK	320
QY	301	ENSTDYLYPEQLKMTVKLISHRCQPPHYVGSVETTKMLCAADPOWKTDSCQDGGGGL	360
DB	321	ENSTDYLYPEQLKMTVKLISHRCQPPHYVGSVETTKMLCAADPOWKTDSCQDGGGGL	380
QY	361	VCSLQGRMTLTGIVSWGRGKALCKDKPGVYTRVSHFLPWIRSHTKKEENGLAL	411
DB	381	VCSLQGRMTLTGIVSWGRGKALCKDKPGVYTRVSHFLPWIRSHTKKEENGLAL	431

RESULT 2  
 UKBAY  
 u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon  
 C;Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)  
 C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 18-Jun-1999  
 C;Accession: S14687; S08651  
 R;Au, Y.P.T.; Wang, T.W.; Clowes, A.W.  
 Nucleic Acids Res. 18, 3411, 1990  
 A;Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasmin  
 A;Reference number: S14687; MUID:90287734; PMID:2113276  
 A;Accession: S14687  
 A;Molecule type: mRNA  
 A;Residues: 1-433 <AU>  
 A;Cross-references: EMBL:X51935; NID:938130; PIDN:CAA36200.1; PID:g38131  
 C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t  
 C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F;1-20/Domain: signal sequence #status predicted <SIG>  
 F;21-176/Product: plasminogen activator chain A #status predicted <ACH>  
 F;30-61/Domain: EGF homology <EGF>  
 F;69-150/Domain: kringle homology <KRG>  
 F;178-433/Product: plasminogen activator chain B #status predicted <BCH>  
 F;178-421/Domain: trypsin homology <TRY>  
 F;167-298, 208-224, 215-287, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted  
 F;223, 274, 378/Active site: His, Asp, Ser #status predicted  
 F;324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	93.4%;	Score 2148;	DB 1;	Length 433;	
Best Local Similarity	92.5%;	Pred. No. 3	6e-158;		
Matches	383;	Conservative	17;	Mismatches	10;
				Indels	4;
				Gaps	2;
QY	1	SNELHQPNSDCCLNGTGVSNKYFNIHWCNCPKFGGQHCIDSKSKTCYEGNGHGYRG	60		
DB	21	SREL-QVPSDCGLNGTGVSNKYFSSIHWCNCPKFGGQHCIDSKSKTCYEGNGHGYRG	79		
QY	61	KASTDTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNCRNPNRRPWCYVOVGLK	120		
DB	80	KASTDTMGRSCLAWNATVLQOYTHAHRSDALQGLGKHNCRNPNRRPWCYVOVGLK	139		
QY	121	PLVQECWHDCAKPKSPPEBELKFCQCGQKTLRPRFKIIGGFTTIENQFWFAAIYRRH	180		
DB	140	QRVOECWHDCAKPKSPPEBELKFCQCGQKTLRPRFKIIGGFTTIENQFWFAAIYRRH	199		
QY	181	RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNNTQGMKFEVENLI	240		
DB	200	RGSVTVYVCGSLISPCWVVSATHCFINYPKEDYIVLGRSLNNTQGMKFEVENLI	259		
QY	241	LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQ---FGTSCBITG	297		
DB	260	LHEDYSADTLAHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPDPFGTSCBITG	319		
QY	298	FGKENTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDCQGDG	357		
DB	320	FGKENTDYLYPEQLKMTVVKLVSHQKCOQPHYGVSEVTTKMLCAADPOWKTDCQGDG	379		
QY	358	GPLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHTEENGLAL	411		
DB	380	GPLVCSIQGHMTLTGIVSGRGCAKDKPGVYTRVSRFLPWIRSHTEENGLAL	433		
RESULT 3					
UNKEG					
u-plasminogen activator (EC 3.4.21.73) precursor - pig					
N:Alternate names: uPA					
C:Species: Sus scrofa domestica (domestic pig)					
C>Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998					
C:Accession: A00932					
R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, B.					
Nucleic Acids Res. 12, 9525-9541, 1984					
A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.					
A:Reference number: A00932; MUID:85087954; PMID:6096832					
A:Accession: A00932					
A:Molecule type: DNA					
A:Residues: 1-240,'H',242-442 <NA>1					
A:Experimental source: kidney cell line LLC-PK1					
R:Nagamine, Y.					
submitted to the Protein Sequence Database, December 1986					
A:Reference number: A37566					
A:Contents: annotation; correction to residue 241					
C:Genetics:					
A:Introns: 19/3, 31/1, 67/1, 125/2, 165/1, 238/2, 288/1, 335/1, 384/3					
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try					
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase					
F:1-20/Domain: signal sequence #status predicted <SIG>					
F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>					
F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>					
F:72-153/Domain: EGF homology <EGF>					
F:72-153/Domain: kringle homology <KRG>					
F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>					
F:190-430/Domain: trypsin homology <TRY>					
F:152/Binding site: carbohydrate (Asn) (covalent) #status predicted					
F:179-310,220-236,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted					
F:235,286,387/Active site: His, Asp, Ser #status predicted					
Query Match					
Best Local Similarity					
Matches					
334; Conservative					
33; Mismatches					
43; Indels					
11; Gaps					
QY	1	SNELHQPNSDCCLNGTGVSNKYFNIHWCNCPKFGGQHCIDSKSKTCYEGNGHGYRG	58		

Db	21	SHELHQPNSDCCLNGGKCVSKYFNSIQRCSPKKFOGHECIDTSQTCPEGNHGS	80		
QY	59	RKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNCRNPNRRPWCYVOVG	118		
DB	81	RKANTNTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNCRNPNRRPWCYVOVG	140		
QY	119	LKPLVQECWHDCA-----DGKPKSPPEBELKFCQCGQKTLRPRFKIIGGFTTIEN	169		
DB	141	LKQVQECWHDCA-----DGKPKSPPEBELKFCQCGQKTLRPRFKIIGGFTTIEN	200		
QY	170	QPMFAAIYRRHSGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNNTQ	229		
DB	201	QPMFAAIYRRHSGSVTVYVCGSLISPCWVVSATHCFINYPKEDYIVLGRSLNNTQ	260		
QY	230	GEMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQ	289		
DB	261	GEMKFEVENLILHEDYSADTLAHNDIALKIRSKGRCQAQPSRTIOTICLPSMYNDPQ	320		
QY	290	GTSCBITGFGKENTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPO	349		
DB	321	GTSCBITGFGKEDPSDYLYPEQLKMTVVKLVSHRECOQPHYGVSEVTTKMLCAADPO	380		
QY	350	DSQCGSDGGLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHTEENGL	409		
DB	381	DSQCGSDGGLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSRFLPWIRSHTEENGL	440		
QY	410	A 410			
DB	441	A 441			
RESULT 4					
JN0560					
u-plasminogen activator (EC 3.4.21.73) precursor - bovine					
N:Alternate names: uPA					
C:Species: Bos primigenius taurus (cattle)					
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999					
C:Accession: JN0560					
R:Kraatzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.					
Gene 125, 177-183, 1993					
A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and indi					
A:Reference number: JN0560; MUID:93216119; PMID:8385052					
A:Accession: JN0560					
A:Molecule type: mRNA					
A:Residues: 1-433 <KRA>					
A:Cross-references: GB:I03546; NID:G163800; PIDN:AAA51419.1; PID:G163801					
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t					
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase					
F:1-20/Domain: signal sequence #status predicted <SIG>					
F:21-179/Product: plasminogen activator chain A #status predicted <MA1>					
F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <MA1>					
F:33-64/Domain: EGF homology <EGF>					
F:72-153/Domain: kringle homology <KRG>					
F:181-433/Product: plasminogen activator					
F:181-421/Domain: trypsin homology <TRY>					
F:170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted					
F:226,277,378/Active site: His, Asp, Ser #status predicted					
Query Match					
Best Local Similarity					
Matches					
310; Conservative					
45; Mismatches					
56; Indels					
2; Gaps					
QY	1	SNELHQPNSDCCLNGTGVSNKYFNIHWCNCPKFGGQHCIDSKSKTCYEGNGHGYRG	58		
DB	21	SNEVHKGESGCGCLNGKCVSKYFNSIQRCSPKKFOGHECIDTSQTCPEGNHGS	80		
QY	59	RKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNCRNPNRRPWCYVOVG	118		
DB	81	RKANDLSGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNCRNPNRRPWCYVOVG	140		
QY	119	LKPLVQECWHDCAKPKSPPEBELKFCQCGQKTLRPRFKIIGGFTTIENQFWFAAIYR	178		
DB	141	LKQVQECWHDCAKPKSPPEBELKFCQCGQKTLRPRFKIIGGFTTIENQFWFAAIYR	200		



Db 150 HDCSLKPKSSVDQFCQKALRPRFKIVGGEFTVENQPFPAIYQKNGKSPSP 209  
Qy 188 VCGSLISPCWVWISATHCFIDYPKKEDYIVYLGSRRLNSNTQGMKFEVENLILHKDYSA 247  
Db 210 KCGSLISPCWVWISATHCFIDYPKKEDYIVYLGSRRLNSNTQGMKFEVENLILHKDYSA 247  
Qy 248 DTLAHHNDIALKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKNSDYL 307  
Db 270 DSLAYHNDIALKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKNSDYL 329  
Qy 308 YPEQLKMTVVKLISHRECOQPHYYGSEVTTKMLCAADPQWKTDCQGDGGLVCSLQGR 367  
Db 330 YPKNLKMSVVKLVSHRECOQPHYYGSEVTTKMLCAADPQWKTDCQGDGGLVCSLQGR 367  
Qy 368 MTLTGIVSWGRCALKDKPGVYTRVSHPLWIRSHTEENGLA 410  
Db 390 PTLGIVSWGRCALKDKPGVYTRVSHPLWIRSHTEENGLA 432

RESULT 7  
A35005  
u-plasminogen activator (EC 3.4.21.73) precursor - chicken  
N;Alternate names: uPA  
C;Species: Gallus gallus (chicken)  
C;Date: 20-Jul-1990 #sequence\_revision 20-Jul-1990 #text\_change 16-Jul-1999  
C;Accession: A35005  
R;Leslie, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.  
J. Biol. Chem. 265, 1339-1344, 1990  
A;Title: The chicken urokinase-type plasminogen activator gene.  
A;Reference number: A35005; PMID:90110185; PMID:2295632  
A;Accession: A35005  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-434 <LS>  
A;Cross-references: GB:J05187; NID:g212858; PIDN:AAA49131.1; PID:g212859  
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
F;40-71/Domain: EGF homology <EGF>  
F;79-158/Domain: kringle homology <KRG>  
F;173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
F;173-416/Domain: trypsin homology <TRY>  
F;162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted  
F;217,272,373/Active site: His, Asp, Ser #status predicted

Query Match 43.6%; Score 1004; DB 1; Length 434;  
Best Local Similarity 47.4%; Pred. No. 6.6e-70;  
Matches 188; Conservative 64; Mismatches 123; Indels 22; Gaps 6;

Qy 11 CDCLINGTCVSNKYFSNIHWCNCPKFGQHCIEDSKTCYEGNGHFYRGKASTDTWGRP 70  
Db 40 CQCLNGTCTIYRFFSQIKCLCEGFGGLHCEIDTNSICYSGNGEDYRGMAEDP----G 95  
Qy 71 CLPWNSATVLQ-QTYHAHRSDALQGLGKHNYCRNPNRRRPWCYVQWGLKPLVQECMVH 129  
Db 96 CLYWDHPSVIRWGDYHADLNALQGLGKHNYCRNPNRRRPWCYVQWGLKPLVQECMVH 149  
Qy 130 DCADGKPPSPPELKCQCKTLRPRFKIIGEFTEINQPFPAIYRHRGSGVTVVC 189  
Db 150 -----TPCSTIEKCECTQSFQSKYFKIVGGSQAQVETQPMIAGIFQNM-GTDQFLC 202  
Qy 190 GGLISPCWVWISATHCFIDYPKKEDYIVYLGSRRLNSNTQGMKFEVENLILHKDY 245  
Db 203 GGLSIDPCWVWISATHCFIDYPKKEDYIVYLGSRRLNSNTQGMKFEVENLILHKDY 245  
Qy 246 SADTLAHHNDIALKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKNSD 305  
Db 263 TDHTGGNDNDIALIRINTASGQCAVESYVVRTCLPEKNLNDYNTWCEIAGYQKNSYD 322  
Qy 306 YLYPEQLKMTVVKLISHRECOQPHYYGSEVTTKMLCAADPQWKTDCQGDGGLVCSLQ 365

Db 323 IYYAQLMSATVNLISQDDCKNKYYDSTRVTDNMVACGDLWETDACKGDSGPMVCEHN 382  
Qy 366 GRMTLTGIVSWGRCALKDKPGVYTRVSHPLWIRSH 402  
Db 383 GRMTLYGIVSWGRCALKDKPGVYTRVSHPLWIRSH 419

RESULT 8  
JS0598  
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat  
N;Alternate names: tissue plasminogen activator  
C;Species: Desmodus rotundus (common vampire bat)  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C;Accession: JS0598  
R;Kraetschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D  
Gene 105, 229-237, 1991  
A;Title: The plasminogen activator family from the salivary gland of the vampire bat D  
A;Reference number: JS0597; PMID:92039036; PMID:1937019  
A;Accession: JS0598  
A;Molecule type: mRNA  
A;Residues: 1-477 <KRA>  
A;Cross-references: GB:M63988; NID:g166074; PIDN:AAA31593.1; PID:g166075  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-36/Domain: propeptide #status predicted <PRO>  
F;37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>  
F;42-79/Domain: fibronectin type I repeat homology <1FA>  
F;87-120/Domain: EGF homology <EGF>  
F;128-209/Domain: kringle homology <KRG>  
F;226-471/Domain: trypsin homology <TRY>  
F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359  
F;185,398/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;225-226/Cleavage site: His-Ser (plasmin) #status predicted  
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 37.6%; Score 864.5; DB 2; Length 477;  
Best Local Similarity 42.9%; Pred. No. 4.2e-59;  
Matches 178; Conservative 60; Mismatches 148; Indels 29; Gaps 9;

Qy 3 ELHQVP----SNCDCLNGTCTVSNKYFSNIHWCNCPKFGQHCIEDSKTCYEGNGHFY 58  
Db 78 QCHTVPVKSCSELRCFNGTCTVSNKYFSNIHWCNCPKFGQHCIEDSKTCYEGNGHFY 136  
Qy 59 RGKASTDTMGRPCLPWNSATVLQTYHAHRSDALQGLGKHNYCRNPNRRRPWCYVQVG 118  
Db 137 RGTWSTSEGAQCINWNNSLLTRTYNGRRSDAITLGLGNHNYCRNPNRRRPWCYVQVG 196  
Qy 119 LKPLVQECMVHDCADGKPPPEELKFCQG-QKTLRPRFKIIGEFTEINQPFPAIY 177  
Db 197 SKFILEFCSPVPCS-----KATGLRKYKEPQLHSTGGLFTDITSHPWQAIF 244  
Qy 178 RHRGGS-VTVVCGGSLISPCWVWISATHCFIDYPKKEDYIVYLGSRRLNSNTQGMKFE 235  
Db 245 AQNRSSGERFLCGGILISSCWLTAAHCFQRYPPHPLVV-LGRYRVKPGKEOTFE 303  
Qy 236 VENLILHKDYADTLAHHNDIALKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCEI 295  
Db 304 VERKIVHEEFDIT--YNNDIALLQLKSGSPQCAQESDVRALCLPEANLQLPDWTECEL 361  
Qy 296 TGFGENSTDYLYPEQLKMTVVKLISHRECOQPHYYGSEVTTKMLCAADPQWKT----- 349  
Db 362 SGYKHKSSPFFSEQLKEGHVLPSSRCTSKFLNKTNNMLCAGDTRSGEIPNVH 421  
Qy 350 DSCQGDGGLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHPLWIRSHTK 404  
Db 422 DACQGDGGLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHPLWIRSHTK 476

RESULT 9  
UKHUT  
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human  
N;Alternate names: t-PA; tissue plasminogen activator

C;Species: Homo sapiens (man)  
C;Date: 14-Nov-1993 #sequence\_revision 14-Nov-1993 #text\_change 08-Dec-2000  
C;Accession: A94004; A23529; J0562; A93293; S02125; A91343; A93951; A91322; A54645; I60  
R;Ny, T.; Elgh, F.; Lund, B.  
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984  
A;Title: The structure of the human tissue-type plasminogen activator gene: correlation  
A;Reference number: A94004; MUID:84298137; PMID:6089198  
A;Accession: A94004  
A;Molecule type: DNA  
A;Residues: 1-562 <NVT>  
A;Cross-references: GB:L00141  
A;Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translation  
R;Frieznher Degen, S.J.; Rajput, B.; Reich, E.  
J. Biol. Chem. 261, 6972-6985, 1986  
A;Title: The human tissue plasminogen activator gene.  
A;Reference number: A23529; MUID:86196143; PMID:3009482  
A;Accession: A23529  
A;Molecule type: DNA  
A;Residues: 1-562 <DEG>  
A;Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818  
R;Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuuda, S.; Higashio, K.  
Agric. Biol. Chem. 55, 1225-1232, 1991  
A;Title: Purification and characterization of tissue plasminogen activator secreted by h  
A;Reference number: J0562; MUID:91291340; PMID:1368681  
A;Accession: J0562  
A;Molecule type: mRNA  
A;Residues: 31-562 <ITA>  
A;Cross-references: DBJ:J001096; NID:G220128; PIDN:BAA00881.1; PID:G441174  
A;Experimental source: embryonic lung fibroblast IMR-90 cells  
A;Note: part of this sequence, including the amino end of the mature protein, was confir  
R;Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, G.A.; Ward, C.A.; Bennett  
Nature 301, 214-221, 1983  
A;Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche  
A;Reference number: A93293; MUID:83115262; PMID:6337343  
A;Accession: A93293  
A;Molecule type: mRNA  
A;Residues: 1-562 <PEN>  
A;Cross-references: GB:L00141  
A;Experimental source: melanoma cells  
R;Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.  
Nucleic Acids Res. 16, 5695, 1988  
A;Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe  
A;Reference number: S02125; MUID:88262579; PMID:3133640  
A;Accession: S02125  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-562 <SAS>  
A;Cross-references: EMBL:X07393; NID:G37243; PIDN:CAA30302.1; PID:G37244  
A;Experimental source: fetal lung cells  
R;Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma  
FEBS Lett. 189, 145-149, 1985  
A;Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen  
A;Reference number: A91343; MUID:85285620; PMID:3896853  
A;Accession: A91343  
A;Molecule type: mRNA  
A;Residues: 1-38, 'G', 86-433, 'E', 435-562 <KAG>  
A;Experimental source: Detroit 562 cells; ATCC 138  
R;Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.  
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983  
A;Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activ  
A;Reference number: A93951; MUID:83169656; PMID:6572897  
A;Accession: A93951  
A;Molecule type: mRNA  
A;Residues: 251-358 <EDL>  
A;Experimental source: melanoma cells  
R;Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.  
Biochemistry 23, 3701-3707, 1984  
A;Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am  
differences.  
A;Reference number: A90488; MUID:85000468; PMID:6433976  
A;Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and  
R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.  
FEBS Lett. 168, 29-32, 1984

A;Title: Differences between uterine and melanoma forms of tissue plasminogen activator  
A;Reference number: A91322; MUID:84158956; PMID:6538514  
A;Accession: A91322  
A;Molecule type: protein  
A;Residues: 33-45;311-320 <POH>  
A;Experimental source: uterus  
A;Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln  
R;van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.  
J. Biol. Chem. 261, 14214-14218, 1986  
A;Reference number: A37567; MUID:87033611; PMID:3021732  
A;Contents: annotation; fibrin binding site  
R;Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Eng  
EMBO J. 5, 3525-3530, 1986  
A;Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen  
A;Contents: annotation; fibrin binding site  
R;Dodd, I.; Nunn, B.; Robinson, J.H.  
Thromb. Haemost. 59, 523-528, 1988  
A;Title: Isolation, identification and pharmacokinetic properties of human tissue-type  
A;Reference number: A60902; MUID:89044681; PMID:3142086  
A;Contents: annotation; novel forms of expressed recombinant t-PA  
R;Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opdenakker, G.  
Mol. Biol. Med. 3, 279-292, 1986  
A;Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its ex  
A;Reference number: A54645; MUID:86284200; PMID:3090401  
A;Accession: A54645  
A;Molecule type: mRNA  
A;Residues: 1-562 <HAR>  
A;Cross-references: GB:M15518; NID:G190031; PIDN:AAA60111.1; PID:G190032  
A;Note: parts of this sequence were confirmed by peptide sequencing  
R;Reddy, V.B.; Garramone, A.J.; Sasaki, H.; Wei, C.  
DNA 6, 461-472, 1987  
A;Title: Expression of human uterine tissue-type plasminogen activator in mouse cells  
A;Reference number: I60110; MUID:88054470; PMID:2824147  
A;Accession: I60110  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-562 <RES>  
A;Cross-references: GB:M18182; NID:G340176; PIDN:AAA36800.1; PID:G340177  
R;Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.  
J. Biol. Chem. 260, 11223-11230, 1985  
A;Title: Isolation and characterization of the human tissue-type plasminogen activator  
A;Reference number: I55232; MUID:85289338; PMID:3161893  
A;Accession: I55232  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-36 <RE2>  
A;Cross-references: GB:M11890; NID:G339837; PIDN:AAA61213.1; PID:G339839  
C;Comment: Cleavage by plasmin or trypsin produces two chains held together by a singl  
C;Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond.  
C;Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat  
C;Genetics: GDB:PLAT  
A;Gene: GDB:119496; OMIM:173370  
A;Cross-references: 8p12-8p12  
A;Map position: 8p12-8p12  
A;Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 5  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase  
F1-23/Domain: signal sequence #status predicted <SIG>  
F1-23/Domain: propeptide #status predicted <PRO>  
F1-23/Domain: signal sequence #status predicted <SIG>  
F33-310/Product: t-plasminogen activator #status experimental <NAT>  
F33-310/Product: t-plasminogen activator chain A #status experimental <ACH>  
F41-78/Domain: fibronectin type I repeat homology <IFI>  
F41-78/Domain: EGF homology <EGF>  
F46-119/Domain: kringle homology <KR1>  
F127-208/Domain: kringle homology <KR2>  
F215-296/Domain: kringle homology <KR2>  
F311-562/Product: t-plasminogen activator chain B #status experimental <BCH>  
F311-566/Domain: trypsin homology <TRY>  
F41-71-69-78-86-97-91-108-110-119-127-208-148-190-179-203-215-236-236-278-291-299  
F152-483/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F129/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental  
F310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental  
F357-406/Active site: His, Asp #status predicted

F:513/Active site: Ser #status experimental

Query Match 37.6%; Score 864.5; DB 1; Length 562;  
Best Local Similarity 38.1%; Pred. No. 5.1e-59;  
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;

QY 3 ELHQVP-----SNCDCLNGGTCVSNKYFNSIHWNCNPKKFGQHCETDKSKTCYEGNGHFY 58  
DB 78 QCHTVFKVKSCELRFCFNGGTCWQAASFSDF-VQCQPKGYTGKQCEVDTHATCYADQGVTY 136  
QY 59 RGKASDTWGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRRRCWCVQVG 118  
DB 137 RGTWSTSEGAQCINWNSNLLTERTYNGRRSDAILGLGNHNYCRNPDRNNSKPCVYKA 196  
QY 119 LKPLVQECMVHDCADGKPPSPPEELKFCQG-QKTLRPRFKIIGGEFTTIENQWFAAIY 177  
DB 197 SKFILEFCSPVPCS-----KATCGLRKYKEPQLHSTGGGLFTDITSHPWQAIF 244  
QY 178 RHRRGS-VTVYCGSLISPCWVISATHCFID-YPKKEDVIVYVGLSRSLNSNTQGEKFE 235  
DB 245 AQNRSSGRFLCGGILISSCWVLTAAHCFQERYPPQHLRV-LGRYRVKPGKEQTPE 303  
QY 236 VENLIILHKYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPOFGTSCFI 295  
DB 304 VEKCIVHEPDDDT--YNDIALQLKSGSPQCAQESDSVRAICLPEANLQLPDWTECEL 361  
QY 296 TGRGENSTDYLPQELKMTVVKLISHREBQQPHYGVSEVTTKMLCAADPQWKT----- 349  
DB 362 SGYGHKSSSPFYSEQLKEGHVRLPSSRCTSKFLNKTVTKNMLCAGDTRSGEIHENVH 421  
QY 350 DSCQSGGGLVCSLQGRMTLTGIVSGRGCALDKPGVTVRVSHPLPWRSHTK 404  
DB 422 DACQSGGGLVCRNDNHMTLLGLISWVGCGEKDIPGVTVKTVNYLGMWRDNR 476

## RESULT 10

A34369  
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)  
C;Species: Megaderma lyra  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A34369  
R;Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob  
J. Biol. Chem. 264, 17947-17952, 1989  
A;Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin  
A;Reference number: A34369; MUID:90036867; PMID:2509450  
A;Accession: A34369  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-477 <GAR>  
A;Cross-references: GB:J05082; NID:g166080; PIDN:AAA31596.1; PID:g166081  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-36/Domain: propeptide #status predicted <PRO>  
F;37-477/Product: plasminogen activator #status predicted <PLA>  
F;42-79/Domain: fibronectin type I repeat homology <IFA>  
F;87-120/Domain: EGF homology <EGF>  
F;128-209/Domain: kringle homology <KRG>  
F;226-471/Domain: trypsin homology <TRY>  
F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4  
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 37.5%; Score 863.5; DB 1; Length 477;  
Best Local Similarity 42.9%; Pred. No. 5.1e-59;  
Matches 178; Conservative 60; Mismatches 148; Indels 29; Gaps 9;

## RESULT 11

JS0599  
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat  
N;Alternate names: tissue plasminogen activator  
C;Species: Desmodus rotundus (common vampire bat)  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C;Accession: JS0599  
R;Kraeuschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D  
Gene 105, 229-237, 1991  
A;Title: The plasminogen activator family from the salivary gland of the vampire bat D  
A;Reference number: JS0597; MUID:92039036; PMID:1937019  
A;Accession: JS0599  
A;Molecule type: mRNA  
A;Residues: 1-431 <KRA>  
A;Cross-references: GB:M63989; NID:g166076; PIDN:AAA31594.1; PID:g166077  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-36/Domain: propeptide #status predicted <PRO>  
F;37-431/Product: plasminogen activator beta #status predicted <PLA>  
F;41-74/Domain: EGF homology <EGF>  
F;82-163/Domain: kringle homology <KRG>  
F;180-425/Domain: trypsin homology <TRY>  
F;41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide b  
F;139,352/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;179-180/Cleavage site: His-Ser (plasmin) #status predicted  
F;226,275,382/Active site: His, Asp, Ser #status predicted  
F;345-361,378-406/Disulfide bonds: #status predicted

Query Match 37.3%; Score 858.5; DB 2; Length 431;  
Best Local Similarity 43.2%; Pred. No. 1.1e-58;  
Matches 175; Conservative 59; Mismatches 146; Indels 25; Gaps 8;

QY 9 SNCDCLNGGTCVSNKYFNSIHWNCNPKKFGQHCETDKSKTCYEGNGHFYRGKASTTMG 68  
DB 42 SELRCFNGGTCWQAASFSDF-VQCQPKGYTGKQCEVDTHATCYADQGVTVRGVSTSESG 100  
QY 69 RPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRRRCWCVQVGLKPLVQECMV 128  
DB 101 AQCNWNSNLLTERTYNGRRSDAILGLGNHNYCRNPDRNNSKPCVYKAKSFILFCSV 160  
QY 129 HDCADGKKPSPPEELAFQCG-QKTLRPRFKIIGGEFTTIENQWFAAIYRRHRGGS-VT 186



```
Db 161 PVCS-----KATCGLRKYPQLHSTGGGLFTDITSHPWQAALPAQNRSSGER 208
QY 187 YVCGSLISPCWVISATHCFID-YPKKEDYIVVLGRSLNSNTQGMKFEVENLIHKKDY 245
Db 209 FICGGILISCCWLTAHCFQERYFPQHLRVV-LGRTYRVKPKQEQTEVEKCIIEEF 267
QY 246 SADTLAHNDIALLKIRSEKGRCAQSRRTIQTCLFSMYNDPQFGSCSEITGKGNSTD 305
Db 268 DDDT--YNNDIALLQLKSGSPQCAQESDSVRAICLPEANLQLPDWTECELSGYGKHKSS 325
QY 306 YLYPEQLKMTVVKLISHRECCQPHYYGSEVTTKMLCAADPQWKT-----DSCQGSQGP 359
Db 326 PFYSEQLKEGHVLYPSSRCTSKLFNKVTNNMLCAGDTRSGEIPNVHVDACQGSQGP 385
QY 360 LVCSLQGRMTLTGIVSGRCALKDKPGVYTRYVSHFLPWIRSHTK 404
Db 386 LVCWNDNHMTLLGIISWVGCGEKDIPGVYTKVNYLIGWIRDNR 430

RESULT 12
A35029
t-plasminogen activator (EC 3.4.21.68) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A35029; A31597
J;Feng, P.; Ohlsson, M.; Ny, T.
J. Biol. Chem. 265, 2022-2027, 1990
A;Title: The structure of the TATA-less rat tissue-type plasminogen activator gene. Spec
A;Reference number: A35029; PMID:90130448; PMID:2105315
A;Accession: A35029
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-559 <FEN>
A;Cross-references: GB:M31197; NID:G207429; PIDN:AAA42261.1; PID:G207431; GB:J05226
R;Ny, T.; Leonardsson, G.; Hsueh, A.J.W.
DNA 7, 671-677, 1988
A;Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator
A;Reference number: A31597; PMID:89170114; PMID:3148445
A;Accession: A31597
A;Molecule type: mRNA
A;Residues: 1-379, 'K', 381-559 <NYT>
A;Cross-references: GB:M23697; NID:G530159; PIDN:AAA41812.1; PID:G530160
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-29/Domain: propeptide #status predicted <PRO>
F;30-559/Product: t-plasminogen activator #status predicted <MAT>
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F;38-75/Domain: fibronectin type I repeat homology <1FI>
F;83-116/Domain: EGF homology <EGF>
F;124-205/Domain: kringle homology <KR1>
F;213-294/Domain: kringle homology <KR2>
F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F;353-553/Domain: trypsin homology <TRY>
F;38-68, 66-75, 83-94, 88-105, 107-116, 124-205, 145-187, 176-200, 213-294, 234-276, 265-289, 297-4
F;149, 481/Binding site: carboxydrate (Asn) (covalent) #status predicted
F;308-309/Cleavage site: Arg-116 (plasmin, trypsin) #status predicted
F;355, 404, 510/Active site: His, Asp, Ser #status predicted

Query Match 36.7%; Score 844.5; DB 1; Length 559;
Best Local Similarity 36.9%; Pred. No. 1.8e-57;
Matches 182; Conservative 64; Mismatches 150; Indels 97; Gaps 12;

QY 3 ELHQP-----SNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHYF 58
Db 74 QCHSVPRVSRSEPRCFNGGTCQALYFSDF-VCQCPDGFVGKRCIDIDTRATCFEGQITY 132
QY 59 RGKASTDTMGPCLPWNSSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYQVG 118
Db 133 RGTWSTAENGACINWNSALSQRPYSARRPNAIKLGLGNHNYCRNPNDRDPKPCYFKA 192
QY 119 LKPLVQECM-----VHDCADGK-----KPSPP----- 141
```

```
Db 193 GKYTEFCSTPACPCKPTEDCTYVGKGVTYRGTHSFTTSKASCLPWNMILGKTYTAWRA 252
QY 142 -----BEKFQ-----COQKTLR-PRFKIIG 162
Db 253 NSQALGLGRHNYCRNPDGDAKPMCHVMKDKLTIWEYCDMSPCSTCGLRQYKQPFRIKGG 312
QY 163 EFTTIENQWFAAIY-RRHGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVVLGR 221
Db 313 LFTDITSHPWQRAIFVKNKRSPGERFLCGVLSSCWLSSAAHCFVERFPFHHLKVVVLGR 372
QY 222 SRLNSNTQGMKFEVENLIHKKDYSADTLAHNDIALLKIRSEKGRCAQSRRTIQTICLP 281
Db 373 TYRVVGEQEOTEIKYIVHKEFDDT--YNDIALQLRSSQCAQESSVGTACL 430
QY 282 SMYNDPQF---CTSCETITGFGKENSTDYLPQKMTVVKLISHRECCQPHYYGSEVTT 337
Db 431 ---DPDVLQPDWTECELSGYKHEASSPFPSDRLEAHVRLYPSRCTSHLFNKTITS 486
QY 338 KMLCAADP-----QWKTDSQDGGSLYCSLOGRMTLTGIVSGRCALKDKPGVYTRY 392
Db 487 NMLCAGDTRTGGNQDHDACQGSQGPLVCMIDKRMTLGLIISWGLCGCGQKDPFGIYTKV 546
QY 393 SHFLPWIRSHTK 405
Db 547 TNYLNIQDNMKQ 559

RESULT 13
JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0597
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidel, W.; Bringmann, P.; Alagon, A.; D
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat D
A;Reference number: JS0597; PMID:92039036; PMID:1937019
A;Accession: JS0597
A;Molecule type: mRNA
A;Residues: 1-477 <KRA>
A;Cross-references: GB:M63987; NID:G166070; PIDN:AAA31591.1; PID:G166071
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
F;42-79/Domain: fibronectin type I repeat homology <1FA>
F;87-120/Domain: EGF homology <EGF>
F;128-209/Domain: kringle homology <KRG>
F;226-471/Domain: trypsin homology <TRY>
F;42-72, 70-79, 87-98, 92-109, 111-120, 128-209, 149-191, 180-204, 214-345, 257-273, 265-334, 359
F;153, 398/Binding site: carboxydrate (Asn) (covalent) #status predicted
F;225-226/Cleavage site: His-Ser (plasmin) #status predicted
F;272, 321, 428/Active site: His, Asp, Ser #status predicted

Query Match 36.4%; Score 836.5; DB 2; Length 477;
Best Local Similarity 42.2%; Pred. No. 6.1e-57;
Matches 175; Conservative 56; Mismatches 147; Indels 37; Gaps 10;

QY 5 HQVPSN-CD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHYRG 60
Db 80 HTVFNVNSCSPRCFNGGTCQWAVYFSDF-VCQCPAGTYGKRCVDTTRATCYEGGVYTRG 138
QY 61 KASTDTMGPCLPWNSSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYQVGLK 120
Db 139 TWSTAERSEVCINWNSLLTRTYNGRMPDAFNGLGNHNYCRNPNGAPKPCYVIKAGK 198
QY 121 PLVQECMVHDCADGKPKSPPEBELKFCQG-QKTLRPRFKIIGGFTTIENQWFAAIYRR 179
Db 199 FTSESCSVFVCS-----KATCGLRKYPQLHSTGGGLFTDITSHPWQAALFAQ 246
QY 180 HRGGS-VTVYVCGSLISPCWVISATHCFIDYPKKEDY-----VYLGRSLNSNTQGMK 233
```



```

Db 247 NRSSGERFLCGGILLISSCWMLTAACF-----QSYLPDLQKVLGRITVRVPGEEQT 301
QY 234 FEVENILHKDYADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPPQFTSC 293
Db 302 FKVKYIVHKEFDDT--YNDIALQLKSDSPQCAQESVRAICLPEANLQLPDWTEC 359
QY 294 EITGFKENSTLYPEQLKMTVVKLISHRECQOPHYGVSEVTKMLCAADPOWKT----- 349
Db 360 ELGSGYGHKSSSPFYEQLEKGVHVRVLPSSRCAPKFLFNKTVTNMMLCAGDTSGEIYPN 419
QY 350 --DSCGDSGGPLVCSLQGRMTLTGIVSWGRGCAKDKKPGVTVRVSHFLPWIRSH 402
Db 420 VHDACQDSGGPLVCMNDNHTLLGIISWGVGCKDVPGVYTKVNYLGIWRDN 474

RESULT 14
A29941
t-plasminogen activator (EC 3.4.21.68) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A29941; S48205; S48207; S48206
R:Rickles, R.J.; Darrow, A.L.; Strickland, S.
J. Biol. Chem. 263, 1563-1569, 1988
A:Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA
A:Reference number: A29941; MUID:88087303; PMID:2826484
A:Accession: A29941
A:Molecule type: mRNA
A:Residues: 1-559 <RIC>
A:Cross-references: GB:J03520; NID:g202109; PION:AAA40470.1; PID:g202110
R:Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A:Title: Characterization of the murine plasma fibrinolytic system.
A:Reference number: S48202; MUID:95010076; PMID:7523120
A:Accession: S48205
A:Molecule type: protein
A:Residues: 33-37, 'X', 39-40 <LIJ>
A:Accession: S48207
A:Molecule type: protein
A:Residues: 309-316 <LIJ>
A:Accession: S48206
A:Molecule type: protein
A:Residues: 33-37, 'X', 39-40 <LIW>
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-23/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <1fl>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KR1>
F:213-294/Domain: kringle homology <KR2>
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: trypsin homology <TRY>
F:38-66, 66-75, 83-94, 88-105, 107-116, 124-205, 145-187, 176-200, 213-294, 234-276, 265-289, 297-4
F:149, 481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-ile (plasmin, trypsin) #status predicted
F:355, 404, 510/Active site: His, Asp, Ser #status predicted

Query Match 36.1%; Score 831.5; DB 1; Length 559;
Best Local Similarity 35.9%; Pred. No. 1.8e-56;
Matches 177; Conservative 65; Mismatches 154; Indels 97; Gaps 10;

QY 3 ELHQVP-----SNCDLNGTGVSNKYFSNTHWNCPCPKKFGQGHCEIDKSKTCVEGNHGFY 58
Db 74 QCHSVPRVSRSEPRCFNGTGCQQAIFYSDF--VCQCPDGFVGKRCIDTRATCFEEQITY 132
QY 59 RGRASDTMTGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDNRRRRCWYQVG 118
Db 133 RGTWSTAESGAECINMNSVLSUKPYNARRPNAIKLGLGNHNYCRNPDRLKPKWCYVFK 192
QY 119 LKPLVQECMVHDCADGKKPS----- 138

```

```

Db 193 GKTTTECSTPACPKGSEDCYVGKGYTYRGTHSLTTSQASCLPWNISVLMGKSYTAWRT 252
QY 139 -----SPPEELK-----FQCG-QKTLRRPRFKTIIG 162
Db 253 NSQALGLARHNYCRNPDGDARPWCHVKDKRKLWETWCMDSPCSTCGLRQYKRFQFRIKGG 312
QY 163 EFTTIENQWFAIY-RRHRGSGVTVVCGSLISPCWVISATHCFIDYPKKEDIYVVLGR 221
Db 313 LYTDITSHPMQAAIFVKNKRSPGFRFLCGGLVLISSCWVLSAAHCFLEFRFPNHLKVVLR 372
QY 222 SRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLP 281
Db 373 TYRVVGESEOTEIEKYIVHEEFDDET--YNDIALQLKRSQKCAQESSVGTACLP 430
QY 282 SMYNDPQF-----GTSCEITGFGKENSTLYLPBQLKMTVVKLISHRECQOPHYGVSEVTT 337
Db 431 ----DENLQLPDWTECELSGYKHEASSPPFDRLKEAHVRLYFPSSRCTSHLFNKTVTN 486
QY 338 KMLCAADP-----QWKTDSCGDSGGPLVCSLQGRMTLTGIVSWGRGCAKDKKPGVTVRV 392
Db 487 NMLCAGDTRSGNODLHDACQDSGGPLVCMINKQMTLTGIIISWGLCGCGKQDVPGVYTKV 546
QY 393 SHELPMIRSHSTKE 405
Db 547 TNYLDWIHNMKKQ 559

RESULT 15
JS0600
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0600
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat D
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0600
A:Molecule type: mRNA
A:Residues: 1-394 <KRA>
A:Cross-references: GB:M63990; NID:gl66078; PIDN:AAA31595.1; PID:gl66079
A:Note: the authors translated the codon ATC for residue 75 as Thr
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-394/Product: plasminogen activator gamma #status predicted <PLA>
F:45-126/Domain: kringle homology <KRG>
F:143-388/Domain: trypsin homology <TRY>
F:45-126, 66-108, 97-121, 131-262, 174-190, 182-251, 276-351, 308-324, 341-369/Disulfide bonds
F:142-143/Cleavage site: His-Ser (plasmin) #status predicted
F:189, 238, 345/Active site: His, Asp, Ser #status predicted
F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 32.7%; Score 752; DB 2; Length 394;
Best Local Similarity 42.0%; Pred. No. 1.6e-50;
Matches 155; Conservative 56; Mismatches 134; Indels 24; Gaps 7;

QY 45 DKSKTCYEGNGHFRYGKASTDTMTGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRN 104
Db 40 DPHATCYKQGVTVRGTWSTESGAQCINWNSNLLIRTYNGMPDEAVRGLGNHNYCRN 99
QY 105 PDNRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEELKFCQG-QKTLRRPRFKTIIG 163
Db 100 PDGASKEPCWIVIRAKFTSSCSVPVCS-----KATCGLRKYKEPQLHSTGGL 147
QY 164 FTTIENQWFAIYRHRGGS-VTVVCGSLISPCWVISATHCFID-YPKKEDIYVVLGR 221
Db 148 FTDITSHPMQAAIFVKNKRSPGFRFLCGGLVLISSCWVLSAAHCFQBRYPQLHVRV-LGR 206
QY 222 SRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLP 281

```

```
Db 207 TYRKPGEQTEVERKCIVHEEEDDT--YNNDIALLQLKSGSPQCAQESDSVRAICLP 264
Qy 282 SMYNDPQFGTSCETITGKKNSTDYLPQQLKMTVVKLISHRECCQPHYSGEYTTKMLC 341
Db 265 EANIQLPDWTECELSGKGKSSFFSEQLKEGHVLYSSRSTKSFELFKNTVNNMLC 324
Qy 342 AADPQWKT-----DSQSGSGPLVCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHF 395
Db 325 AGDTRSGEIVPNVHDAQQSGSGPLVCMNDNHTLLGLIISGVGCGEKDIPGVYTKVTNY 384
Qy 396 LPWIRSHTK 404
Db 385 LGWIRDNR 393

RESULT 16
A46688
hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
C:Accession: A46688
R: Miyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J. Biol. Chem. 268, 10024-10028, 1993
A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease
d coagulation factor XII.
A:Reference number: A46688; MUID:93252878; PMID:7683665
A:Accession: A46688
A:Molecule type: mRNA
A:Residues: 1-655 <MY>
A:Cross-references: DBJ: D14012; NID: g219680; PID: BAA03113.1; PID: g219681
A:Experimental source: liver (mRNA); serum (protein)
A:Note: sequence extracted from NCBI backbone (NCBI: 131227, NCBI: 131228)
A:Note: parts of the sequence, including the amino ends of the heavy and light chains, c
C:Genetics:
A:Gene: GDB: HGPAC; HGPA; HGPAP
A:Cross-references: GDB: 9954514
A:Map position: 4p16-4p16
C:Function:
A:Description: activates hepatocyte growth factor by specific proteolytic cleavage
A:Pathway: tissue repair and regeneration
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
F:1-34/Domain: signal sequence #status predicted <SIG>
F:108-148/Domain: fibronectin type II repeat homology <IF2>
F:164-197/Domain: EGF homology <EG1>
F:202-237/Domain: fibronectin type I repeat homology <IF1>
F:245-278/Domain: EGF homology <EG2>
F:286-367/Domain: kringle homology <KRG>
F:373-407/Product: hepatocyte growth factor activator light chain #status experimental <
F:408-655/Product: hepatocyte growth factor activator heavy chain #status experimental <
F:408-641/Domain: trypsin homology <TRY>
F:40,48,290,469,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:164-175,169-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-36
F:447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 32.0%; Score 735.5; DB 1; Length 655;
Best Local Similarity 37.6%; Pred. No. 5.4e-49;
Matches 160; Conservative 53; Mismatches 166; Indels 47; Gaps 9;

Qy 5 HQVPSNCDCLNGGTCVSNKYFSNIHW-----CNCPKFGQGHCEIDKSKTCYEGNGH 56
Db 242 HTACLSFCLNGGTC-----HLIVATGTVTACPPGAGRLCNLEPDERCFLNGT 292
Qy 57 FYRKASDTWGRPCLPNWSATVLOQTHAHSRALQLGLGKHNYCRNPNDRRPPWCYVQ 116
Db 293 GYRGVASTSAGSLSCLANWSDLLQELHVDVSGAALLGLGPHAYCRNPNDRPWCYV 352
Qy 117 VGLPVLQECMVHDC-----ADGKPSPEELKFCQGK-----TLRPRFX 158
Db 353 KDSALSWCYKLEACESLTRVQLSPDLLATLPEPASGRQ---ACGRHKKRTFLRPR-- 407
Qy 159 IIGGEFTTIQNPFAAIYRRHRGSGVTYVCGGSLISPCWVISATHCFIDYPKKEDYIV 218
```

```
Db 408 IIGSSSLPGSHPLAAIY---IGDS---FCAGSLVHTCMWVSAHCFSPPRDSVSW 461
Qy 219 LGRSLRNSQTGEMKFEVENLIHKDYSADTLAHNDIALKIRSKSGRCACQPSRTTQTI 278
Db 462 LGQHFRRTDVTQTFOIEKYIPTLYSVNPSDH-DLVLRLLKKGDRCATRQFQVPI 520
Qy 279 CLPSMYNDPQFGTSCETITGKKNSTDYLPQQLKMTVVKLISHRECCQPHYSGEYTTK 338
Db 521 CLPBPGSTFPAGHKCQIAGWGHLDENVSGYSSSLREALVPLVADHKCSSPEVYGADISPN 580
Qy 339 MLCRAADPQWKTDCQSGSGPLVCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFPLW 398
Db 581 MLCAGYFDCKSDACQSGSGPLACRKNGVAYLIIISWGDCGRLHKGVYTRVANYVDW 640
Qy 399 IRSHTK 404
Db 641 INDIR 646

RESULT 17
S28941
coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)
N: Alternate names: Hageman factor
C: Species: Cavia porcellus (guinea pig)
C: Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000
C: Accession: S28941
R: Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.; Okabe,
Biochim. Biophys. Acta 1159, 113-121, 1992
A: Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage
A: Reference number: S28941; MUID: 93003367; PMID: 1390917
A: Accession: S28941
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-603 <SM>
A: Cross-references: EMBL: X68615; NID: g49578; PID: CAA48600.1; PID: g49579
A: Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology
C: Keywords: hydrolase; serine proteinase
F: 46-87/Domain: fibronectin type II repeat homology <IF2>
F: 134-169/Domain: fibronectin type I repeat homology <FB1>
F: 177-208/Domain: EGF homology <EGF>
F: 216-294/Domain: kringle homology <KRG>
F: 359-597/Domain: trypsin homology <TRY>

Query Match 31.2%; Score 718.5; DB 2; Length 603;
Best Local Similarity 36.8%; Pred. No. 1e-47;
Matches 161; Conservative 64; Mismatches 149; Indels 63; Gaps 13;

Qy 13 CLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFGYRGKASTDTMGRPCL 72
Db 182 CLNGGRCLE---VEGHLCDLCPMGYTGPFCDLDTTASCYEGRGVSYRGMAETTVSGAKCQ 238
Qy 73 PWSATVLOQTHAHSRD-ALQLGLGKHNYCRNPNDRRPPWCYVQVGLKPLVQECMVHDC 131
Db 239 RWAS---EATYRNWTAQALRGLGHTTCRNFNDNDRTPWCFVWMGNLSWEYCDLAQC 294
Qy 132 ADGKPSPEELKFO-----CGQKTLRPR 156
Db 295 QYPPQPTATPHD-RFEHPKLPSSRLSILQTPQPTONQALANELPETSSLLCGOR-LRKR 352
Qy 157 F---KIIGGEFTTIQNPFAAIYRRHRGSGVTYVCGGSLISPCWVISATHCFIDYPKK 212
Db 353 LSSLSRIVGLVLPAGHPYIADY-----WGS--NFCGSLIAPCWLTIAHCLQNRPAP 406
Qy 213 EDYIVYLRGRSLRNSQTGEMKFEVENLIHKDYSADTLAHNDIALKIRSKSGRCACQPSRTTQTI 271
Db 407 EELKVLGQDRHNSQCEHCQTLAVHSYRLHEAFSPS--SYLNDLALLRLQKSDAGSCAQL 464
Qy 272 SRTIQTCLPSMYNDPQFG--TSCEITGFGKKNSTDYLPQQLKMTVVKLISHRECCQPH 329
Db 465 SPYVQTVCLPSGAPPSESETTCCEVAGWGQHFEGAEYSFSLQEAQVPLISSRCSPE 524
Qy 330 YGSEVITKMLCAADPQWKTDCQSGSGPLVCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFPLW 398
```



C;Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MATL>  
F;75-106/Domain: EGF homology <EG1>  
F;113-145/Domain: EGF homology <EG2>  
F;152-185/Domain: EGF homology <EG3>  
F;192-274/Domain: kringle homology <KRI>  
F;312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MATs>  
F;312-548/Domain: trypsin homology <TRY>

Query Match 29.3%; Score 674.5; DB 2; Length 558;  
Best Local Similarity 38.0%; Pred. No. 2.3e-44;  
Matches 159; Conservative 56; Mismatches 164; Indels 37; Gaps 16;

QY 13 CLNGGTCVSNKYFSNIHWNCNPKKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMGPRCL 72  
DB 157 CQNGGVCSSRRHRSRF--TCACPDQYKGFCEIGPD--DCYVGDGYSYRGKSVKTVNQNPCL 214  
QY 73 PWSNATVLOQTYHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLKPLVQE--CMVHDC 131  
DB 215 YWNSHLLQTYNMFEMEDAETHGIAEHNFRCNPDGDHKPWCFCVKNSEKVKWEYCDYTV 274  
QY 132 ADGKKPSPPEEL-----KFO--CGQKTLRPRF--KIIGGEFTTIENQFWFAIY--- 177  
DB 275 PVPDTP--NPVESLLEPVMELPFCFESCGKTEVAEHAVKRIYGGFKSTAGKHPWQVLSQ 333  
QY 178 ----RRHRGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIYVLSRSLNSNTQGMK 233  
DB 334 PLATSMFQGG---HFCGALIHPCWLVTAHC--TDINTKHLKV--LGDQDLKKTESHEQT 387  
QY 234 FEVENLILHKDYSADTLAHNDIALKIRSEKGRCAQPSRTIOTICLPSMYNDPQF--GT 291  
DB 388 FRVEKILKYQYNERDEIPHNDIALKLPVGGHCALESRYKVTVCPLP---DFFPSGTE 444  
QY 293 CEITGFGKENSTDYLYPEQLKMTVVKLIHRECCQPHYGGSEVTTKMLCAADPOWK--TDS 351  
DB 445 CHISGWGVTGE--GSRQLLDAKVKLIANPLCNSRQLYDHTIDDSMICAGNLQKPGSDT 502  
QY 352 CQGDSSGGLVCSLQGRMTLTGIVSWRGCAKDKPGVYTRVSHFLPWIRSHKTEENGL 409  
DB 503 CQGDSSGGLTCEKDGTYGVYVIGVSWGRCG--KKPGVYTVQVTKFLNWKITMTMREAGL 558

## RESULT 20

JC4795  
plasma hyaluronan-binding protein precursor - human  
N;Alternate names: hepatocyte growth factor activator-like protein; PHBP  
N;Contains: serine proteinase (EC 3.4.21.-)  
C;Species: Homo sapiens (man)  
C;Date: 15-Oct-1995 #sequence\_revision 16-Aug-1996 #text\_change 19-Jul-2002  
C;Accession: JC4795  
R;Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tomita, M.  
J. Biochem. 119, 1157-1165, 1996  
A;Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP)  
r activator.

A;Reference number: JC4795; MUID:96425001; PMID:8827452  
A;Accession: JC4795  
A;Molecule type: mRNA  
A;Residues: 1-560 <CHO>  
A;Cross-references: GB:S83182; NID:g1836158; PIDN:AAB46909.1; PID:g1836159  
A;Experimental source: plasma  
A;Note: parts of this sequence, including the amino ends of the mature chains, were deter  
C;Genetics:  
A;Gene: GDB:HABP2; HABP; PHBP; HGFA1  
A;Cross-references: GDB:4573962  
C;Complex: a disulfide-bonded heterodimer of chains produced from the same precursor; th  
C;Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin  
C;Keywords: chondroitin sulfate proteoglycan; glycoprotein; hyaluronic acid; hydrolase;  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-313/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>  
F;77-108/Domain: EGF homology <EG1>  
F;115-147/Domain: EGF homology <EG2>  
F;154-187/Domain: EGF homology <EG3>

F;194-276/Domain: kringle homology <KRI>  
F;314-550/Domain: trypsin homology <TRY>  
F;314-516/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted  
F;54-207/Binding site: carbohydrate (Asn)  
F;77-88,82-97,99-108,115-125,120-136,138-147,154-165,159-176,178-187,194-276,215-257,24  
F;362,405/509/Active site: His, Asp, Ser #status predicted

Query Match 28.7%; Score 661.5; DB 1; Length 560;  
Best Local Similarity 37.1%; Pred. No. 2.3e-43;  
Matches 155; Conservative 58; Mismatches 166; Indels 39; Gaps 15;

QY 13 CLNGGTCVSNKYFSNIHWNCNPKKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMGPRCL 72  
DB 159 CQNGATCSRHRKRSKF--TCACPDQFGKFCBIG--SDDCYVGDGYSYRGKWNRTVNOHACL 216  
QY 73 PWSNATVLOQTYHAHRSDALQLGLGKHNCRNPNRRPWCYVQVGLKPL-----VQEC 126  
DB 217 YWNSHLLQTYNMFEMEDAETHGIEHNFRCNPDADKPCFKIVTNDKVKWEYCDVSAC 276  
QY 127 MVHDC--DGKKPSPPEELK--FQCGQKTLRPR--FKIIGGEFTTIENQFWFAIYRRHR 181  
DB 277 SAQDVAVPEESPTEPSTKLPGFDSCGKTEAERKIRIYGGFKSTAGKHPWQASLQ---- 332  
QY 182 GGSVT-----YVCGSLISPCWVISATHCFIDYPKKEDYIYVLSRSLNSNTQGMK 233  
DB 333 --SSLPLTISMPQGHFCGALIHPCWLVTAHC--TDI--KTRHLKVVLGGDQLKKEEFHEQS 389  
QY 234 FEVENLILHKDYSADTLAHNDIALKIRSEKGRCAQPSRTIOTICLPSMYNDPQF--GT 291  
DB 390 FRVEKIFKYSHYNERDEIPHNDIALKLPVGGHCALESKYVKTVCPLP---DGSFSGS 445  
QY 292 SCEITGFGKENSTDYLYPEQLKMTVVKLIHRECCQPHYGGSEVTTKMLCAADPOWK--TD 350  
DB 446 ECHISGWGVTETGK--GSRQLLDAKVKLIANTLCSRLQYLDHIDDSMICAGNLQKPGQD 503  
QY 351 SCQDSSGGLVCSLQGRMTLTGIVSWRGCAKDKPGVYTRVSHFLPWIRSHKTEENG 408  
DB 504 TCQDSSGGLTCEKDGTYGVYVIGVSWGLECG--KKPGVYTVQVTKFLNWKITATIKESG 559

## RESULT 21

S45281  
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)  
N;Alternate names: Hageman factor (activated)  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 10-Apr-1995 #sequence\_revision 22-Apr-1995 #text\_change 21-Jan-2000  
C;Accession: S45281; A61329  
R;Shibuya, Y.; Samba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.  
Biochim. Biophys. Acta 1206, 63-70, 1994  
A;Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): cor  
A;Reference number: S45281; MUID:94242782; PMID:8186251  
A;Accession: S45281  
A;Molecule type: mRNA  
A;Residues: 1-593 <SHI>  
A;Cross-references: GB:S70164  
A;Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70  
is, and ATC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as (C  
R;Fujikawa, K.; Walsh, K.A.; Davie, E.W.  
Biochemistry 16, 2270-2278, 1977  
A;Title: Isolation and characterization of bovine factor XII (Hageman factor).  
A;Reference number: A61329; MUID:77182112; PMID:861210  
A;Accession: A61329  
A;Molecule type: protein  
A;Residues: 10-16,'X',18-19;525-550 <FUG>  
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homolog  
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;  
F;37-78/Domain: fibronectin type II repeat homology <IF2>  
F;88-120/Domain: EGF homology <EGF>  
F;115-160/Domain: fibronectin type I repeat homology <FB1>  
F;207-287/Domain: kringle homology <KR>  
F;350-587/Domain: trypsin homology <TRY>  
F;541/Active site: Ser #status predicted

```
Query Match          27.9%; Score 642; DB 2; Length 593;
Best Local Similarity 35.1%; Pred. No. 7.9e-42;
Matches 155; Conservative 60; Mismatches 167; Indels 60; Gaps 15;

QY 6 QVPSNCDCILNGTGVSNKYFSNIHWCNCPKFGQHCEDKSKTCYE--GNGHFYRGKAS 63
DB 166 QVCRINPCLPWSATVLOQTY-HAHRSDALQGLGHKHNCRPNDRRRPWCYVQVGLKPL 122
QY 64 TDTMGRPCLPWSATVLOQTY-HAHRSDALQGLGHKHNCRPNDRRRPWCYVQVGLKPL 122
DB 223 TTLSGAPCQSWAS---EATYNNVTAEQVLNWLGLGDHAFCRPNPDNTRPWCFTWKGDRLS 278
QY 123 VQECMVHDC--ADG-----KPSS-----PPEELKF-----QCQG 150
DB 279 WNYCRLPACQAAGHEHFLPLSPSALQKPESTTQTPLPSLTSGWCSTFTPLASGPGCGCQ 338
QY 151 ---KTLRPRFKIIGGFTTIENQPFALVRRHGGSVTVVCGGSLISPCWVISAATHCFI 207
DB 339 RLKWLSSLNRVGVLVALPGHPIYALYWDQ-----HFCAGSLIATPCWULTAAHCLQ 392
QY 208 DVPKXEDYIVYLGSRSLNNTQGMKFEVENLILHKDYSADTLAHNDIALKIR-SKRG 266
DB 393 NRPAPKELTWVLQDRHQNSCEQCQTLAVRDYRLHEAFSPITYQH--DLALVRLQESADG 450
QY 267 RQAQSRITQITCLPSMYNDPQFGTS--CEITGPKENSTDYLYPQLKMTVVKLISHRE 324
DB 451 CCAHPSFPVQVCLPSTAARPAEAAVCEVAGWHQFEGGE--YSSFLEAQVPLIDPQR 509
QY 325 CQOPHYGVSEVTTKMLCAADPOWKTDSCQDSSGGLVLC---SLQGRMTLTGIVSWRGCA 381
DB 510 CSAPDVHGAFTQGLMCAGLEGGTDACQDSSGGLVLCVDETPERQLILRGIVSWSGCG 569
QY 382 LKDKPGVYTRVSHFLPWIRSH 403
DB 570 NRLKPGVYTDVANYLAWIREHT 591

RESULT 22
B61545
plasmin (EC 3.4.21.7) precursor - sheep (fragments)
N;Alternate names: plasminogen
N;Contains: miniplasminogen
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
C;Accession: B61545; S28200
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A;Title: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; PMID:89005015; PMID:3168975
A;Accession: B61545
A;Molecule type: protein
A;Residues: 1-37;38-117 <SCH>
R;Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 5, 21-25, 1992
A;Title: Complete amino acid sequence of ovine miniplasminogen.
A;Reference number: S28200; PMID:93149995; PMID:1492092
A;Accession: S28200
A;Molecule type: protein
A;Residues: 118-460 <SC2>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase; z
F;1-37,38-117,118-460/product; plasminogen (fragments) #status experimental <PRO>
F;1-37/Domin: activation peptide (fragment) #status experimental <APT>
F;38-117,118-230,231-460/product; plasmin (fragments) #status experimental <MAT>
F;41-118/Domin: kringle homology <KR4>
F;118-460/Product: miniplasminogen #status experimental <MIN>
F;132-211/Domin: kringle homology <KR5>
F;226-460/Domin: plasmin chain B #status experimental <BCH>
F;231-453/Domin: trypsin homology <TRI>
F;272,315,410/Active site: His, Asp, Ser #status predicted

Query Match          22.1%; Score 507.5; DB 2; Length 460;
```

```
Best Local Similarity 30.1%; Pred. No. 1.4e-31;
Matches 139; Conservative 51; Mismatches 147; Indels 125; Gaps 15;

QY 35 KKFGGG-----HCEIDSKTCYEGNHFYRGKASTDTMGRPCLPWSATV--LQOTYH 85
DB 20 KKLAGEVDECAKCE-EBAQDCYHNGGQYRGTSSTTVTGRKCCQSSSMIPHRHQKTP 78
QY 86 AHRSDALQGLGHKHNCRPNDRRRPWCYVQ-----VGLK-----PLVQECMVHDC 132
DB 79 SYPNAGLTW-----NYCRNPDAKSPWCYTTPRVRWEFCNLKKAQAPSVPBPADCM 133
QY 133 -----DCKPSS-----139
DB 134 LGIGKYRGKATTVAGVPCQEWAAQEPHRHGIFTPETNPRAGLEKXNYCRNPGDVNGPW 193
QY 140 -----PPEELKFCQGOKTLRPR---FKIIGGEFTTIENQPFALVRRHR 181
DB 194 CYTTPNPKLFDYCIDPQCESSFDGPKVPKPCPARVVGCVATPHSWPQVSLRRSR 253
QY 182 GGSVTVVCGGSLISPCWVISAATHCFIDYPKKEDYIVYLGSRSLNNTQGMKFEVENLIL 241
DB 254 -----EHFCGGLTISPEWVLTAAHCLDSILGPSFTYVILGAHYEMAREASVQEI 309
QY 242 HKDYSADTLAHNDIALKIRSKEGRCAQPSRTIQTICLPSMYNDPQF---GTSCEITG 297
DB 310 EFSRA-----DIALKLSSP-----AVITDEVIPACLPS-----PNYVVADKTVCYITG 353
QY 298 FGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDSSG 357
DB 354 WGETQGT--FGVGRLEKEARLPVLENKVCNRYEVLNGRVKSSTELCAGDLAGGTDSCQDSSG 411
QY 358 GPLVCSLQGRMTLTGIVSWRGRCALKDKPGVYTRVSHFLPWI 399
DB 412 GPLVCFEKDKYILQGVTSWGLGCARDPKPGVYRVYSTYVPMI 453
```

## RESULT 23

```
PLFG
plasmin (EC 3.4.21.7) precursor - pig (fragment)
N;Alternate names: plasminogen
N;Contains: miniplasminogen
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 07-Sep-1990 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C;Accession: S03733; S03737; A25834
R;Schaller, J.; Marti, T.; Roesele, S.J.; Kaempfer, U.; Rickli, E.E.
Fibrinolysis 1, 91-102, 1987
A;Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the c
A;Reference number: S03733
A;Accession: S03733
A;Molecule type: protein
A;Residues: 1-560 <SCH>
R;Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,
A;Reference number: S03735; PMID:81212097; PMID:7238497
A;Accession: S03737
A;Molecule type: protein
A;Residues: 1-57 <BRU>
R;Marti, T.; Schaller, J.; Rickli, E.E.
Eur. J. Biochem. 149, 279-285, 1985
A;Title: Determination of the complete amino-acid sequence of porcine miniplasminogen.
A;Reference number: A25834; PMID:85203907; PMID:3846533
A;Accession: A25834
A;Molecule type: protein
A;Residues: 450-790 <MAR>
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a \
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen ac
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homoloq
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma; serine prot
F;1-790/Product: plasminogen #status predicted <PRO>
F;1-77/Domin: plasminogen-related protein precursor homology (fragment) <PLPH>
```

F:1-77/Domain: activation peptide #status predicted <APT>  
F:78-560/Product: plasmin chain A #status predicted <ACH>  
F:84-162/Domain: kringie homology <KR1>  
F:166-243/Domain: kringie homology <KR2>  
F:256-333/Domain: kringie homology <KR3>  
F:358-435/Domain: kringie homology <KR4>  
F:450-790/Product: miniplasminogen #status experimental <MIN>  
F:461-540/Domain: kringie homology <KR5>  
F:561-780/Product: plasmin chain B #status experimental <BCH>  
F:561-783/Domain: trypsin homology <TRY>  
F:330-54,34-42,84-162,105-145,133-157,166-243,169-297,187-226,215-238,256-333,277-316,305  
bonds: #status predicted  
F:602,645,740/Active site: His, Asp, Ser #status predicted

Query Match 21.9%; Score 504.5; DB 1; Length 790;  
Best Local Similarity 35.4%; Pred. No. 4.4e-31;  
Matches 130; Conservative 45; Mismatches 141; Indels 51; Gaps 14;

Qy 45 DSKITCVGNGHFRGKASTDTMGPCLPWNSATVLOQTYAHR-----SDALQLGLGKHN 100  
Db 456 DLSEDCMFGNGKRYGRKRAITVAGVPCQEWAA-----QEPHRSIFTPETNPRAGLEK-N 509

Qy 101 YCRNPD-NRRPPWCVVQGLPLVQECMVHDCADGKXSPPEELKFCGOKTLRPR--- 156  
Db 510 YCRNPDGDDNGPWYIT-TNPQKLFYCDVPCQVTS-----SFDGPKVPEPKCP 558

Qy 157 FKIIIGBFTTIENQFWFAIYRRHGGSVTVVCGSLISPCWISATHCFIDYPKKEDIY 216  
Db 559 ARVVGGCVSIPHSFWQISLRYVRG---HFCGGLTSLPEWVLTAHCKLEKSSPSYK 614

Qy 217 VYLGSRSLNNTQGMKEFEVENLILHKDYSADTLAHNNDIALHKIRSEKGCAPSRITQ 276  
Db 615 VILGAHEYHILGEGVQIEDVSKLF--KEPS-----EADIALKLSSP-----AVITDKVI 662

Qy 277 TICLPSMYNDPQF---GTSCEITGFGKENSTDYLPQLKMTVVKLISHRECQOPHYG 332  
Db 663 PACLPT---PNYVVADRTACYITGWGETKT--YGAGLLKEARLPVLENKVCNRYEVLG 716

Qy 333 SEVTVMKLCADPOWKTSCGSDGGPLVCSLQGRMTLTGTVSGRGCAKDKPGVYTRV 392  
Db 717 GKVSFNLKAGLAGGIDSCGSDGGPLVCFEKDKYIIQGVTSWGLGALPKPKPGVYVRV 776

Qy 393 SHFLPWI 399  
Db 777 SRFVTWI 783

RESULT 24  
PLHU

Plasmin (EC 3.4.21.7) precursor [validated] - human  
N;Alternate names: plasminogen precursor [misnomer]  
N;Contains: angiotatin; microplasmin; plasminogen  
C;Species: Homo sapiens (man)  
C;Date: 24-Apr-1984 #sequence revision 02-Dec-1994 #text change 15-Sep-2000  
C;Accession: A35229; I52242; A26646; I62738; I84626; S03735; A00929; A04627; A04625; A04  
R;Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.  
J. Biol. Chem. 265, 6104-6111, 1990  
A;Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr  
A;Reference number: A35229; MUID:90202879; PMID:2318848  
A;Accession: A35229  
A;Molecule type: DNA  
A;Residues: 1-810 <PEP>  
A;Cross-references: GB:J05286; GB:M34276; NID:g190064; PIDN:AAA60113.1; PID:g387026  
A;Experimental source: leukocyte; lung fibroblast  
R;Malgaretti, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.; Ta  
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990  
A;Title: Definition of the transcription initiation site of human plasminogen gene in li  
A;Reference number: I52242; MUID:91097523; PMID:2266308  
A;Accession: I52242  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-16 <MAL1>  
A;Cross-references: GB:M62890; NID:g190092; PIDN:AAA36454.1; PID:g553613

R;Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.  
FEBS Lett. 213, 254-260, 1987  
A;Title: Molecular cloning and characterization of a full-length cDNA clone for human f  
A;Reference number: A26646; MUID:87162490; PMID:3030813  
A;Accession: A26646  
A;Molecule type: mRNA  
A;Residues: 1-471, 'D', 473-810 <FOR>  
A;Cross-references: GB:X05199; NID:g35530; PIDN:CAA28831.1; PID:g35531  
A;Experimental source: liver  
R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.  
Biochemistry 23, 4243-4250, 1984  
A;Title: Characterization of a complementary deoxyribonucleic acid coding for human an  
A;Reference number: I45961; MUID:85023311; PMID:6148961  
A;Accession: I62738  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 292-471, 'D', 473-810 <MAL2>  
A;Cross-references: GB:K02922; NID:g190112; PIDN:AAA60124.1; PID:g387031  
A;Accession: I84609  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 367-419 <MAL3>  
A;Cross-references: GB:K02921; NID:g190110; PIDN:AAA60123.1; PID:g190111  
R;Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M  
Eur. J. Biochem. 114, 465-470, 1981  
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of human  
A;Reference number: S03735; MUID:81212097; PMID:7238497  
A;Accession: S03735  
A;Molecule type: protein  
A;Residues: 20-71, 'E', 73-76 <BRU>  
R;Sottstrup-Jensen, L.; Petersen, T.E.; Magnusson, S.  
submitted to the Atlas, July 1977  
A;Reference number: A00929  
A;Accession: A00929  
A;Molecule type: protein  
A;Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>  
R;Wiman, B.  
Eur. J. Biochem. 76, 129-137, 1977  
A;Title: Primary structure of the B-chain of human plasmin.  
A;Reference number: A04627; MUID:7725245; PMID:142009  
A;Accession: A04627  
A;Molecule type: protein  
A;Residues: 581-810 <WII>  
R;Wiman, B.; Wallen, P.  
Eur. J. Biochem. 50, 489-494, 1975  
A;Title: Structural relationship between "glutamic acid" and "lysine" forms of human p  
A;Reference number: A04625; MUID:75093329; PMID:122932  
A;Accession: A04625  
A;Molecule type: protein  
A;Residues: 20-50, 'Q', 51-71, 'E', 73-85, 87-100 <WII2>  
R;Wiman, B.; Wallen, P.  
Eur. J. Biochem. 58, 539-547, 1975  
A;Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen t  
A;Reference number: A04626; MUID:76043692; PMID:126863  
A;Accession: A04626  
A;Molecule type: protein  
A;Residues: 483-507, 'E', 509-604 <WII3>  
R;Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.  
J. Biol. Chem. 248, 1631-1633, 1973  
A;Title: The primary structure of human plasminogen. II. The histidine loop of human p  
A;Reference number: A92125; MUID:73145248; PMID:4694729  
A;Contents: annotation; active site  
R;Groskopf, W.R.; Summaria, L.; Robbins, K.C.  
J. Biol. Chem. 244, 3590-3597, 1969  
A;Title: Studies on the active center of human plasmin. Partial amino acid sequence of  
A;Reference number: A92048; MUID:69234739; PMID:4240117  
A;Contents: annotation; active site  
R;Trexler, M.; Vail, Z.; Paddy, L.  
J. Biol. Chem. 257, 7401-7406, 1982  
A;Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen  
A;Reference number: A92382; MUID:82213905; PMID:6919539  
A;Contents: annotation; omega-aminocarboxylic acid binding sites

R,Vali, Z.; Patthy, L.  
 J. Biol. Chem. 259, 13690-13694, 1984  
 A;Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential  
 A;Reference number: A92458; PMID:85054794; PMID:6094526  
 A;Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site  
 J;Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCance, S.G.;  
 J. Biol. Chem. 271, 29461-29467, 1996  
 A;Title: Kringle domains of human angiotensin. Characterization of the anti-proliferative  
 A;Reference number: A58811; PMID:97067211; PMID:8910613  
 A;Contents: annotation  
 R;Lijnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.  
 Biochemistry 37, 4699-4702, 1998  
 A;Title: Generation of an angiotensin-like fragment from plasminogen by stromelysin-1 (M)  
 A;Reference number: A58812; PMID:9548733; PMID:9548733  
 A;Contents: annotation  
 R;Tulinsky, A.; Mulichak, A.M.  
 submitted to the Brookhaven Protein Data Bank, July 1991  
 A;Reference number: A51341; PDB:1PK4  
 A;Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454  
 R;Tulinsky, A.; Wu, T.P.  
 submitted to the Brookhaven Protein Data Bank, August 1993  
 A;Reference number: A51911; PDB:1PKR  
 A;Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181  
 R;Padmanabhan, K.; Tulinsky, A.  
 submitted to the Brookhaven Protein Data Bank, April 1994  
 A;Reference number: A52408; PDB:1PMK  
 A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454  
 R;Tulinsky, A.; Mathews, I.I.  
 submitted to the Brookhaven Protein Data Bank, December 1995  
 A;Reference number: A65244; PDB:1CEA  
 A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181  
 R;Tulinsky, A.; Mathews, I.I.  
 submitted to the Brookhaven Protein Data Bank, December 1995  
 A;Reference number: A65245; PDB:1CEB  
 A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181  
 R;Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.  
 Biochemistry 30, 10576-10588, 1991  
 A;Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å  
 A;Reference number: A58819; PMID:92031502; PMID:1657148  
 A;Contents: annotation  
 R;Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.  
 Biochemistry 30, 10589-10594, 1991  
 A;Title: The refined structure of the epsilon-aminocaproic acid complex of human plasmin  
 A;Reference number: A58818; PMID:92031503; PMID:1657149  
 A;Contents: annotation  
 R;De Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, M.  
 Biochemistry 31, 270-279, 1992  
 A;Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2.1  
 A;Reference number: A39483; PMID:92118803; PMID:1310033  
 A;Contents: annotation; X-ray crystallography, 2.4 angstroms  
 R;Stec, B.; Teeter, M.M.; Whittow, M.; Yamano, A.  
 submitted to the Brookhaven Protein Data Bank, June 1995  
 A;Reference number: A65980; PDB:1KRN  
 A;Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454  
 R;Rejante, M.; Llinas, M.  
 submitted to the Brookhaven Protein Data Bank, August 1996  
 A;Reference number: A65803; PDB:1HPJ  
 A;Contents: annotation; conformation by (1)H-NMR, residues 103-181  
 R;Rejante, M.; Llinas, M.  
 submitted to the Brookhaven Protein Data Bank, August 1996  
 A;Reference number: A65804; PDB:1HPK  
 A;Contents: annotation; conformation by (1)H-NMR, residues 103-181  
 R;Rejante, M.R.; Llinas, M.  
 Eur. J. Biochem. 221, 927-937, 1994  
 A;Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.  
 A;Reference number: S43645; PMID:94237157; PMID:8181475  
 A;Contents: annotation; conformation by (1)H-NMR, residues 96-184  
 R;Rejante, M.R.; Llinas, M.  
 Eur. J. Biochem. 221, 939-949, 1994

A;Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasmin  
 A;Reference number: A58817; PMID:94237158; PMID:8181476  
 A;Contents: annotation; conformation by (1)H-NMR  
 C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many c  
 d PIR:FGHUB).  
 C;Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITHUA2) immediately a  
 rg-580, resulting in two chains connected by two disulfide bonds. Without the inhibitor  
 C;Comment: Microplasmin is formed by autolytic cleavage of plasmin under artificial co  
 C;Comment: Stromelysin 1 (see PIR:KCHUS1) acts on plasminogen to produce angiotensin. 1  
 ting solid tumors.  
 C;Genetics:  
 A;Gene: GDB:PLG  
 A;Cross-references: GDB:119498; OMIM:173350  
 A;Map position: 6q26-6q27  
 A;Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 52  
 C;Function:  
 A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a v  
 us the walls of the graafian follicle; also activates the urokinase-type plasminogen ac  
 C;Pathway: fibrinolysis  
 C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolog  
 C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hys  
 F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
 F;1-19/Domain: signal sequence #status experimental <SIG>  
 F;20-810/Product: plasminogen #status experimental <PRO>  
 F;20-96/Domain: activation peptide #status experimental <APT>  
 F;79-466/Product: angiotensin #status experimental <AST>  
 F;97-580,581-810/Product: plasmin #status experimental <MAT>  
 F;97-580/Domain: plasmin chain A #status experimental <CHA>  
 F;103-181/Domain: kringle homology <KR1>  
 F;185-262/Domain: kringle homology <KR2>  
 F;275-352/Domain: kringle homology <KR3>  
 F;377-454/Domain: kringle homology <KR4>  
 F;481-560/Domain: kringle homology <KR5>  
 F;550-580,581-810/Product: microplasmin #status experimental <MMT>  
 Query Match 21.8%; Score 501; DB 1; Length 810;  
 Best Local Similarity 33.5%; Pred. No. 8.3e-31;  
 Matches 139; Conservative 45; Mismatches 167; Indels 64; Gaps 16;  
 QY 10 NCDCLNGGTCVGNKYFNSIHWNCPCPKFGQ-----HCEIDKSKTCYEGNH 56  
 DB 428 NPDADKGPWCFTDTSVRWEYNLKKCSGTASVAPPPVLLPNVETSEEDCMFGNGK 487  
 QY 57 FYRGKASTDTMGRPCPLMNSATVLOQTYHAHR-----SDALQLGLGKHNYCRNPD-NRRRP 111  
 DB 488 GYRGKRAITVTGTPCQDMAA-----QEPHRHSIFTPTNPRAGLEK-NYCRNPDGDVGGP 541  
 QY 112 WCYVQVGLKPLVQECMVHDCADGKXSPPEBLKQCGQKTLRPR---PKITGGEFTTIE 168  
 DB 542 WCYT-TNPKRLDYCDVPOCA---APS-----FDCGKQPVEPKKCPGRVVGCVAPH 590  
 QY 169 NOPWFAAIYRRHRGSGVTVCVCGSLISPCOWISATHCFIDYPKEDYIVVIGRSLNSNT 228  
 DB 591 SWPQVSL--RTRFG--MHFCGGLISPEWVTAHCKLEKSPRPSYKVLGAHQEWLIE 646  
 QY 229 QEMKFEVENLLHKDYSDATLHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQ 288  
 DB 647 PHVQIEVSRLEPTEPT-----RKDIALLKLSPP---AVITDKVIPACLPS---PN 690  
 QY 289 F----GTSCEITGFGKENSTDYLYPEQLKMTVVVLLKLSHRECCQPHYSSEVTKMLCAAD 344  
 DB 691 YVADRTCEFTIGWGETGT--FGAGLKEAQLPVLENKVCNRYEFLNGRVSOSTELCAGH 748  
 QY 345 POWKTDSQGGSGGGLVCSLQGRMTLTGIVSMGRCALKDKEGVTYTRVSHFLPWI 399  
 DB 749 LAGGTDSCQGGSGGLVCFEKDKYILQGVTSWGLGCARPKNKPGVTVRSRFTWI 803  
 RESULT 25  
 B30848  
 plasmin (EC 3.4.21.7) precursor - rhesus macaque  
 C;Species: Macaca mulatta (rhesus macaque)



R;Tijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.  
Eur. J. Biochem. 224, 863-871, 1994  
A;Title: Characterization of the murine plasma fibrinolytic system.  
A;Reference number: S48202; MUID:95010076; PMID:7523120  
A;Accession: S48202  
A;Molecule type: protein  
A;Residues: 20-25 <LIJ>  
A;Accession: S48203  
A;Molecule type: protein  
A;Residues: 22-27 <LI2>  
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many other tissues. Plasminogen is converted into plasmin by plasminogen activators, both plasma and tissue. Plasmin immediately after dissociation from the clot. In the presence of the inhibitor, the active inhibitor, the activation involves also removal of the activation peptide.  
C;Comment: Stromelysin 1 (see PIR:KCMSS1) acts on plasminogen to produce angiotatin. Tissue plasminogen activator (tPA) acts on plasminogen to produce angiotatin. Tissue plasminogen activator (tPA) acts on plasminogen to produce angiotatin.  
C;Function:  
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of tissues; acts on the walls of the graafian follicle; also activates the urokinase-type plasminogen activator (uPA).  
C;Superfamily: plasmin  
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydrolysis; plasminogen-related protein precursor homology <PLPH>  
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-812/Product: plasminogen #status predicted <PRO>  
F;20-96/Domain: activation peptide #status predicted <APT>  
F;79-466/Product: angiotatin #status predicted <AST>  
F;97-581, 582-812/Product: plasmin #status predicted <MAT>  
F;97-581/Domain: chain A #status predicted <ACH>  
F;103-181/Domain: kringie homology <KR1>  
F;103-181/Domain: kringie homology <KR1>  
F;185-262/Domain: kringie homology <KR2>  
F;275-352/Domain: kringie homology <KR3>  
F;377-454/Domain: kringie homology <KR4>  
F;481-560/Domain: kringie homology <KR5>  
F;582-812/Domain: chain B #status predicted <BCH>  
F;582-805/Domain: trypsin homology <TR1>  
F;49-73, 53-61, 103-181, 124-164, 152-176, 185-262, 188-316, 206-245, 234-257, 275-352, 296-335, 336-373, 385-400, 401-410, 411-420, 421-430, 431-440, 441-450, 451-460, 461-470, 471-480, 481-490, 491-500, 501-510, 511-520, 521-530, 531-540, 541-550, 551-560, 561-570, 571-580, 581-590, 591-600, 601-610, 611-620, 621-630, 631-640, 641-650, 651-660, 661-670, 671-680, 681-690, 691-700, 701-710, 711-720, 721-730, 731-740, 741-750, 751-760, 761-770, 771-780, 781-790, 791-800, 801-810, 811-820, 821-830, 831-840, 841-850, 851-860, 861-870, 871-880, 881-890, 891-900, 901-910, 911-920, 921-930, 931-940, 941-950, 951-960, 961-970, 971-980, 981-990, 991-1000, 1001-1010, 1011-1020, 1021-1030, 1031-1040, 1041-1050, 1051-1060, 1061-1070, 1071-1080, 1081-1090, 1091-1100, 1101-1110, 1111-1120, 1121-1130, 1131-1140, 1141-1150, 1151-1160, 1161-1170, 1171-1180, 1181-1190, 1191-1200, 1201-1210, 1211-1220, 1221-1230, 1231-1240, 1241-1250, 1251-1260, 1261-1270, 1271-1280, 1281-1290, 1291-1300, 1301-1310, 1311-1320, 1321-1330, 1331-1340, 1341-1350, 1351-1360, 1361-1370, 1371-1380, 1381-1390, 1391-1400, 1401-1410, 1411-1420, 1421-1430, 1431-1440, 1441-1450, 1451-1460, 1461-1470, 1471-1480, 1481-1490, 1491-1500, 1501-1510, 1511-1520, 1521-1530, 1531-1540, 1541-1550, 1551-1560, 1561-1570, 1571-1580, 1581-1590, 1591-1600, 1601-1610, 1611-1620, 1621-1630, 1631-1640, 1641-1650, 1651-1660, 1661-1670, 1671-1680, 1681-1690, 1691-1700, 1701-1710, 1711-1720, 1721-1730, 1731-1740, 1741-1750, 1751-1760, 1761-1770, 1771-1780, 1781-1790, 1791-1800, 1801-1810, 1811-1820, 1821-1830, 1831-1840, 1841-1850, 1851-1860, 1861-1870, 1871-1880, 1881-1890, 1891-1900, 1901-1910, 1911-1920, 1921-1930, 1931-1940, 1941-1950, 1951-1960, 1961-1970, 1971-1980, 1981-1990, 1991-2000, 2001-2010, 2011-2020, 2021-2030, 2031-2040, 2041-2050, 2051-2060, 2061-2070, 2071-2080, 2081-2090, 2091-2100, 2101-2110, 2111-2120, 2121-2130, 2131-2140, 2141-2150, 2151-2160, 2161-2170, 2171-2180, 2181-2190, 2191-2200, 2201-2210, 2211-2220, 2221-2230, 2231-2240, 2241-2250, 2251-2260, 2261-2270, 2271-2280, 2281-2290, 2291-2300, 2301-2310, 2311-2320, 2321-2330, 2331-2340, 2341-2350, 2351-2360, 2361-2370, 2371-2380, 2381-2390, 2391-2400, 2401-2410, 2411-2420, 2421-2430, 2431-2440, 2441-2450, 2451-2460, 2461-2470, 2471-2480, 2481-2490, 2491-2500, 2501-2510, 2511-2520, 2521-2530, 2531-2540, 2541-2550, 2551-2560, 2561-2570, 2571-2580, 2581-2590, 2591-2600, 2601-2610, 2611-2620, 2621-2630, 2631-2640, 2641-2650, 2651-2660, 2661-2670, 2671-2680, 2681-2690, 2691-2700, 2701-2710, 2711-2720, 2721-2730, 2731-2740, 2741-2750, 2751-2760, 2761-2770, 2771-2780, 2781-2790, 2791-2800, 2801-2810, 2811-2820, 2821-2830, 2831-2840, 2841-2850, 2851-2860, 2861-2870, 2871-2880, 2881-2890, 2891-2900, 2901-2910, 2911-2920, 2921-2930, 2931-2940, 2941-2950, 2951-2960, 2961-2970, 2971-2980, 2981-2990, 2991-3000, 3001-3010, 3011-3020, 3021-3030, 3031-3040, 3041-3050, 3051-3060, 3061-3070, 3071-3080, 3081-3090, 3091-3100, 3101-3110, 3111-3120, 3121-3130, 3131-3140, 3141-3150, 3151-3160, 3161-3170, 3171-3180, 3181-3190, 3191-3200, 3201-3210, 3211-3220, 3221-3230, 3231-3240, 3241-3250, 3251-3260, 3261-3270, 3271-3280, 3281-3290, 3291-3300, 3301-3310, 3311-3320, 3321-3330, 3331-3340, 3341-3350, 3351-3360,





QY 265 EGRCAQPSRTIQICLIPSMYNDPQF-----GTSCEITGFGKSTDYLYPQLKMTVVKLI 320  
Db 1282 -SRPAITDVKVIFACPLPS-----PNYVITAMTECYITGWTGQT--FGAGLLKEAQLHVI 1334  
QY 321 SHRECCQPHYVYGSEVTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGVSVWGRGC 380  
Db 1335 ENTVCNHYEFLNGRVKSTELCAGHLAGGTDRCCQDNGGPPVCPDKXVILRGITSWGPGC 1394  
QY 381 ALKDKPGVYTRVSHFLPWI 399  
Db 1395 ACNPKPGVYRVSSFTWI 1413  
RESULT 29  
S00657  
apoprotein(a) (EC 3.4.21.-) precursor [validated] - human  
N/Alternate names: apolipoprotein(a); lipoprotein(a) chain apo(a)  
C/Species: Homo sapiens (man)  
C/Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text change 08-Dec-2000  
C/Accession: S00657; A28017; A47277; I60906; A47233; I52415; I65286  
R/McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.; Chen, E.Y.; Fless, G.M.; Scar  
Nature 330, 132-137, 1987  
A/Title: cDNA sequence of human apolipoprotein(a) is homologous to plasminogen.  
A/Reference number: S00657; MUID:88039109; PMID:3670400  
A/Accession: S00657  
A/Molecule type: mRNA  
A/Residues: 1-4548 <MCL>  
A/Cross-references: GB:X06290; EMBL:X06696; NID:G28619; PIDN:CAA29618.1; PID:G28620  
R/Eaton, D.L.; Fless, G.M.; Kohr, W.J.; McLean, J.W.; Xu, Q.T.; Miller, C.G.; Lawn, R.M.  
Proc. Natl. Acad. Sci. U.S.A. 84, 3224-3228, 1987  
A/Title: Partial amino acid sequence of apolipoprotein(a) shows that it is homologous to  
A/Reference number: A28017; MUID:87204109; PMID:3472206  
A/Accession: A28017  
A/Molecule type: protein  
A/Residues: 20-21, 'P', 23-34, 177-179, 'N', 181-186, 'T', 188-196, 'DKG', 200:292-314, 'W', 316-31  
X', 4396-4401 <EAT>  
R/Wade, D.P.; Clarke, J.G.; Lindahl, G.E.; Liu, A.C.; Zysow, B.R.; Meer, K.; Schwartz, K  
Proc. Natl. Acad. Sci. U.S.A. 90, 1369-1373, 1993  
A/Title: 5' control regions of the apolipoprotein(a) gene and members of the related pla  
A/Reference number: A47277; MUID:93165698; PMID:7679504  
A/Accession: A47277  
A/Status: preliminary; translation not shown; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-16 <RES>  
A/Cross-references: GB:I07899; NID:G967973; PID:G967974  
R/Margaret, N.; Acquati, F.; Magnaghi, P.; Bruno, L.; Pontoglio, M.; Rocchi, M.; Sacco  
Proc. Natl. Acad. Sci. U.S.A. 89, 11584-11588, 1992  
A/Title: Characterization by yeast artificial chromosome cloning of the linked apolipop  
A/Reference number: A47233; MUID:93087573; PMID:1454851  
A/Accession: I60906  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-16 <RE2>  
A/Cross-references: GB:M90078; NID:G178786; PIDN:AAA35547.1; PID:G553188  
A/Note: apo(a) gene 1 (nomenclature of reference I52415)  
A/Accession: A47233  
A/Status: preliminary; translation not shown; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-16 <RES>  
A/Cross-references: GB:M90079; NID:G178784; PIDN:AAA35546.1; PID:G553187  
R/Ichinose, A.  
Biochemistry 31, 3113-3118, 1992  
A/Title: Multiple members of the plasminogen-apolipoprotein(a) gene family associated wi  
A/Reference number: I52415; MUID:92207924; PMID:1554698  
A/Accession: I52415  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-16 <RE3>  
A/Cross-references: GB:M86877; NID:G178780; PIDN:AAB49909.1; PID:G553185  
A/Note: apo(a) gene 1 (nomenclature of reference I52415)  
A/Accession: I65286  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA

A/Residues: 1-16 <RE4>  
A/Cross-references: GB:M86878; NID:G178782; PIDN:AAA51749.1; PID:G553186  
C/Genetics:  
A/Gene: GDB:LPA  
A/Cross-references: GDB:120699; OMIM:152200  
A/Map position: 6q26-6q27  
A/Note: several genes closely linked on chromosome 6 are identical in the first coding  
rs of kringle repeats  
C/Superfamily: apolipoprotein(a); kringle homology; trypsin homology  
C/Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase  
F/1-19/Domain: signal sequence #status predicted <SIG>  
F/20-4548/Product: apolipoprotein(a) #status experimental <MAT>  
F/28-105/Domain: kringle homology <KR1>  
F/142-219/Domain: kringle homology <KR2>  
F/256-333/Domain: kringle homology <KR3>  
F/370-447/Domain: kringle homology <KR4>  
F/484-561/Domain: kringle homology <KR5>  
F/598-675/Domain: kringle homology <KR6>  
F/712-789/Domain: kringle homology <KR7>  
F/826-903/Domain: kringle homology <KR8>  
F/940-1017/Domain: kringle homology <KR9>  
F/1054-1131/Domain: kringle homology <KR10>  
F/1168-1245/Domain: kringle homology <KR11>  
F/1282-1359/Domain: kringle homology <KR12>  
F/1396-1473/Domain: kringle homology <KR13>  
F/1510-1587/Domain: kringle homology <KR14>  
F/1624-1701/Domain: kringle homology <KR15>  
F/1738-1815/Domain: kringle homology <KR16>  
F/1852-1929/Domain: kringle homology <KR17>  
F/1966-2043/Domain: kringle homology <KR18>  
F/2080-2157/Domain: kringle homology <KR19>  
F/2194-2271/Domain: kringle homology <KR20>  
F/2308-2385/Domain: kringle homology <KR21>  
F/2422-2499/Domain: kringle homology <KR22>  
F/2536-2613/Domain: kringle homology <KR23>  
F/2650-2727/Domain: kringle homology <KR24>  
F/2764-2841/Domain: kringle homology <KR25>  
F/2878-2955/Domain: kringle homology <KR26>  
F/2992-3069/Domain: kringle homology <KR27>  
F/3106-3183/Domain: kringle homology <KR28>  
F/3220-3297/Domain: kringle homology <KR29>  
F/3334-3411/Domain: kringle homology <KR30>  
F/3448-3525/Domain: kringle homology <KR31>  
F/3562-3639/Domain: kringle homology <KR32>  
F/3676-3753/Domain: kringle homology <KR33>  
F/3782-3859/Domain: kringle homology <KR34>  
F/3896-3973/Domain: kringle homology <KR35>  
F/4010-4087/Domain: kringle homology <KR36>  
F/4124-4201/Domain: kringle homology <KR37>  
F/4228-4307/Domain: kringle homology <KR38>  
F/4328-4541/Domain: trypsin homology <TRY>  
Query Match 20.6%; Score 474; DB 1; Length 4548;  
Best Local Similarity 35.2%; Pred. No. 6, 5e-28;  
Matches 129; Conservative 36; Mismatches 133; Indels 68; Gaps 17;  
QY 50 CYEGNGHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGK-----HNYCRN 104  
Db 4228 CMFGNGKGYRGKKAATVVTGTCQEW-----AAQEPHRRHSTFIPGNTKWAGLEKNYCRN 4280  
QY 105 PD-NRRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEELKFCQCKTLRPR---FKII 160  
Db 4281 PGDINGPWCYT-MNPKULFDYCDIPLCASS-----SFDCKGPQVEPKKCPGSIV 4329  
QY 161 GGEFTTIENQWPFAATYRRHSGSVTVCGGSLISPCWVISATHCFTDYPKCKEDYIVILG 220  
Db 4330 GGCVAHPHSPWQVSL--RTRFGK--HFCGGTILSPDWLTAHCLKKSRRSSSYKVIILG 4385  
QY 221 RRR---LNNTQCEMKFVENILHKDYSADTLAHNHDIALLKIRSKGRCQAPSPRTIQT 277  
Db 4386 AHQEVNLESHVQ---EIEVSRFLFLEPTQA-----DIALLEK-----SRPAVITDKVMP 4430  
QY 278 ICLPSMYNDPQF-----GTSCEITGFGKSTDYLYPQLKMTVVKLI SHRECCQPHYVYG 333

```

Db      4431 ACLEPS-----PYWMTARTECYITWGEGTQGT--FGTGLLKEAQLLVINEVCN--HY--- 4479
Qy      334 EVTTKMLCAADPQWKTDSQQDSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVS 393
Db      4480 ----KYICAEHLARGTDSQQDSGGPLVCFEKKVKYILQGVTSWGLGCARPKNKPGVYARVS 4535
Qy      394 HFLPMI 399
Db      4536 RFVTWI 4541

```

## RESULT 30

```

A61545
Plasmin (EC 3.4.21.7) precursor - horse (fragments)
N:Alternate names: plasminogen
N:Contains: miniplasminogen
C:Species: Equus caballus (domestic horse)
C:Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
R:Accession: A61545; S17527
R:Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A:Title: Structural aspects of the plasminogen of various species.
A:Reference number: A61545; MUID:89005015; PMID:3168975
A:Accession: A61545
A:Molecule type: protein
A:Residues: 1-33,34-117 <SCH>
R:Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 4, 69-74, 1991
A:Title: Complete amino acid sequence of equine miniplasminogen.
A:Reference number: S17527; MUID:92052077; PMID:1946332
A:Accession: S17527
A:Molecule type: protein
A:Residues: 118-455 <SC2>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasmin; serine proteinase; z
F:1-33,34-117,118-455/Product: plasminogen (fragments) #status experimental <PRO>
F:1-33/Domain: activation peptide (fragment) #status experimental <APT>
F:34-117,118-225,226-455/Product: plasmin (fragments) #status experimental <MAT>
F:37-114/Domain: kringle homology <KR4>
F:118-455/Product: miniplasminogen #status experimental <MIN>
F:126-205/Domain: kringle homology <KR5>
F:226-455/Domain: plasmin chain B #status experimental <BCH>
F:226-448/Domain: trypsin homology <TRY>
F:267,310,405/Active site: His, Asp, Ser #status predicted

```

```

Query Match      19.9%; Score 457.5; DB 2; Length 455;
Best Local Similarity 29.5%; Pred. No. 1e-27;
Matches 132; Conservative 45; Mismatches 156; Indels 115; Gaps 15;

```

```

Qy      40 QHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCLPMSATV--LQOTYHAHRSDALQLGLG 97
Db      28 EECEA-KVQDCYQDKGSEYRGTSITVTGKKCQSWSNTPHHQKTPKYPNADLTW--- 83
Qy      98 KHNCRNPDNRPRWCYVQ-----VGLK---PLVQECMVHDC-----DGKK--- 136
Db      84 --NYCRNPDGDKGWCYCTTDPSEVWERFCNLKCKSETVQEPSEPDCLGIGKYGQKKATT 141
Qy      137 -----PSSPP----- 141
Db      142 VTGTRQAAQAQBPHRHSIFTPEANPWANLEKNYCRNPDGDVNGPWCYTMNPQKLFDYCD 201
Qy      142 ----EELKFOGQKTLRPR---FKIIGGEFTTIENQWFAAIYRRHRGGSVTVVCGGSLI 194
Db      202 VPQCESSPFDGKPKFVKPKCSGRIVGGCVIAHSWV--QISLRTFG--RHPCGGTLL 257
Qy      195 SPQWISATCFIDYPKKEDIYVYGLSRSLNSNTQGEKPFVENLIHLKDYSDTLAHN 254
Db      258 SPEWVLTAAHCLERSRPSTYKVVLGTHHELRLAAGAAQIDVSKLFLEPSRA----- 309
Qy      255 DIALLKIRSKEGRCQAQPSRIQTI---CLPSMYNDPOFGTSCEITGPKENSTDYLYPEQ 311
Db      310 DIALLKLSL-----PAITTONVIPACLPADPPVYVANWAECFVTGWGE--TQDSSNAGV 360

```

```

Qy      312 LKMTVVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSQQDSGGPLVCSLQGRMTLT 371
Db      361 LKEAQLPVIENKVCNRXYEYLNGRYKSTELCAGHLVGGVDSQQDSGGPLVCFEKDKYILQ 420
Qy      372 GIVSWGRCALKDKPGVYVTRVSHFLPMI 399
Db      421 GVTWGLGCARPKNKPGVYVTRVSSFINWI 448

```

Search completed: December 3, 2003, 14:44:15  
Job time : 24.2586 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:53 ; Search time 13.2061 Seconds  
(without alignments)  
1463.563 Million cell updates/sec

Title: US-09-880-503-3

Perfect score: 2301

Sequence: 1 SNELHQVPSNCCLNGTGV.....VSHFLPWIRSHYKENGAL 411

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	2298	99.9	431	1	UROK_HUMAN	P00749 homo sapien
2	2148	93.4	433	1	UROK_PAPCY	P16227 papio cynoc
3	1859.5	80.8	442	1	UROK_PIG	P04185 sus scrofa
4	1767	76.8	433	1	UROK_BOVIN	Q05589 bos taurus
5	1690.5	73.5	432	1	UROK_RAT	P29598 rattus norv
6	1660.5	72.2	433	1	UROK_MOUSE	P06869 mus musculu
7	1004	43.6	434	1	UROK_CHICK	P15120 gallus gall
8	864.5	37.6	477	1	URT2_DESRO	P15638 desmodus ro
9	864.5	37.6	562	1	TPA_HUMAN	P00750 homo sapien
10	858.5	37.3	431	1	URTE_DESRO	P98121 desmodus ro
11	844.5	36.7	559	1	TPA_EAT	P19637 rattus norv
12	836.5	36.4	477	1	URTI_DESRO	P98119 desmodus ro
13	826.5	35.9	559	1	TPA_MOUSE	P11214 mus musculu
14	818	35.5	566	1	TPA_BOVIN	Q28198 bos taurus
15	752	32.7	394	1	URTG_DESRO	P49150 desmodus ro
16	735.5	32.0	655	1	HGFA_HUMAN	Q04756 homo sapien
17	726	31.6	653	1	HGFA_MOUSE	Q9r098 mus musculu
18	718.5	31.2	603	1	FAI2_CAVPO	Q04962 cavia porce
19	692	30.1	615	1	FAI2_HUMAN	P00748 homo sapien
20	642	27.9	593	1	FAI2_BOVIN	P98140 bos taurus
21	504.5	21.9	790	1	PLMN_PIG	P06867 sus scrofa
22	501	21.8	812	1	PLMN_MOUSE	P20918 mus musculu
23	499	21.7	810	1	PLMN_HUMAN	P00747 homo sapien
24	497	21.6	810	1	PLMN_MACMU	P12545 macaca mula
25	494.5	21.5	333	1	PLMN_CANFA	P80009 canis famil
26	493	21.4	343	1	PLMN_SHEEP	P81286 ovis aries
27	484.5	21.1	812	1	PLMN_BOVIN	P06868 bos taurus
28	484	21.0	1420	1	APOA_MACMU	P14417 macaca mula
29	474	20.6	4548	1	APOA_HUMAN	P08519 homo sapien
30	462	20.1	875	1	NETR_HUMAN	P56730 homo sapien
31	449.5	19.5	331	1	PLMN_HORSE	P80010 equus cabal
32	445	19.3	761	1	NETR_MOUSE	O08762 mus musculu
33	435	18.9	810	1	PLMN_ERIEU	Q29485 erinaceus e

RESULT 1  
UROK\_HUMAN STANDARD; PRT; 431 AA.  
AC P00749; Q15844; Q16618; Q969W6;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 20-MAR-1987 (Rel. 04, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
DE (U-plasminogen activator).  
GN PLAU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=85215647; PubMed=2987867;  
RX Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;  
RT "The human urokinase-plasminogen activator gene and its promoter.";  
RL Nucleic Acids Res. 13:2759-2771(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,  
RA Steffens G.J., Heyneker H.L.;  
RT "Cloning and expression of the gene for pro-urokinase in Escherichia coli.";  
RL Biotechnology 3:923-929(1985).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86056954; PubMed=2415429;  
RA Nagai M., Hiratsuka R., Kaneda T., Hayasuke N., Arimura H.,  
RA Nishida M., Suyama T.;  
RT "Molecular cloning of cDNA coding for human prepro-urokinase.";  
RL Gene 36:183-188(1985).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85203359; PubMed=3898571;  
RA Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,  
RA van Elsen A., Herzog A., Bollen A.;  
RT "Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA.";  
RL DNA 4:139-146(1985).  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,  
RA Nickerson D.A.;  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

ALIGNMENTS

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahney J., Heiton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[7]  
RN SEQUENCE OF 66-431 FROM N.A.  
RP MEDLINE=84272706; PubMed=6589620;  
RX MEDLINE=83055084; PubMed=6754569;  
RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.;  
RT "Identification and primary sequence of an unspliced human urokinase  
RT poly(A)+ RNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).  
[8]  
RN SEQUENCE OF 21-177.  
RP MEDLINE=83055084; PubMed=6754569;  
RX MEDLINE=83055084; PubMed=6754569;  
RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,  
RA Flohe L.;  
RT "The primary structure of high molecular mass urokinase from human  
RT urine. The complete amino acid sequence of the A chain.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).  
[9]  
RN SEQUENCE OF 156-176 AND 179-224.  
RP MEDLINE=83003608; PubMed=6749491;  
RX Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W.,  
RA Studer R.O.;  
RT "Human low-molecular-weight urinary urokinase. Partial  
RT characterization and preliminary sequence data of the two polypeptide  
RT chains.";  
RL Eur. J. Biochem. 125:251-257(1982).  
[10]  
RN SEQUENCE OF 158-410.  
RP MEDLINE=83055099; PubMed=6754572;  
RX Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;  
RA "The complete amino acid sequence of low molecular mass urokinase  
RT from human urine.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).  
[11]  
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RP MEDLINE=96000858; PubMed=8591045;  
RX Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,  
RA Dobson C.M., Stuart D.I., Jones E.Y.;  
RT "The crystal structure of the catalytic domain of human  
RT urokinase-type plasminogen activator.";  
RL Structure 3:681-691(1995).  
[12]  
RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.  
RP MEDLINE=20266327; PubMed=10805774;  
RX Sperl S., Jacob U., Arroyo de Prada N., Sturzbecher J., Wilhelm O.G.,  
RA Bode W., Magdolen V., Huber R., Moroder L.;  
RT "4-aminomethylphenylguanidine derivatives as nonpeptidic highly  
RT selective inhibitors of human urokinase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).  
[13]  
RN STRUCTURE BY NMR.  
RP MEDLINE=89127526; PubMed=2536903;  
RX Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;  
RA "Dynamics of the multidomain fibrinolytic protein urokinase from two-  
RT dimensional NMR.";  
RL Nature 337:579-582(1989).  
[14]  
RN STRUCTURE BY NMR OF 67-155.  
RX MEDLINE=93003110; PubMed=1327118;  
RA Li X., Smith R.A.G., Dobson C.M.;  
RT "Sequential 1H NMR assignments and secondary structure of the kringle  
RT domain from urokinase.";  
RL Biochemistry 31:9562-9571(1992).  
[15]  
RN STRUCTURE BY NMR OF 67-155.  
RP MEDLINE=94149701; PubMed=8107091;  
RX Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;  
RA "Solution structure of the kringle domain from urokinase-type  
RT plasminogen activator.";  
RL J. Mol. Biol. 235:1548-1559(1994).  
[16]  
RN VARIANT LEU-141.  
RP MEDLINE=96186279; PubMed=8652631;  
RX Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,  
RA Sawasaki Y., Hanada K.;  
RT "Characterization of single chain urokinase-type plasminogen  
RT activator with a novel amino-acid substitution in the kringle  
RT structure.";  
RL Biochim. Biophys. Acta 1293:83-89(1996).  
[17]  
RN VARIANT LEU-141.  
RP MEDLINE=97218551; PubMed=9065988;  
RX Conne B., Berczy M., Belin D.;  
RT "Detection of polymorphisms in the human urokinase-type plasminogen  
RT activator gene.";  
RL Thromb. Haemost. 77:434-435(1997).  
[18]  
RN VARIANT LEU-141.  
RP MEDLINE=97337920; PubMed=9194591;  
RX Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Heil W.,  
RA Creutzburg S., Graeff H., Magdolen V.;  
RT "Mutational analysis of the genes encoding urokinase-type plasminogen  
RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";  
RL Electrophoresis 18:686-689(1997).  
CC -/- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR  
CC THERAPY OF THROMBOLYTIC DISORDERS.  
CC -/- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -/- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS  
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A  
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW  
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.  
CC -/- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used  
CC in Pulmonary Embolism (PE) to initiate fibrinolysis.  
CC -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -/- SIMILARITY: Contains 1 kringle domain.  
CC -/- SIMILARITY: Contains 1 EGF-like domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X02419; CAA26268.1; -  
DR EMBL; M15476; AAA61253.1; -  
DR EMBL; D00244; BAA00175.1; -  
DR EMBL; D11143; BAA01919.1; -  
DR EMBL; X02760; CAA26535.1; -  
DR EMBL; AF377330; AAK53822.1; -  
DR EMBL; BC013575; AAH13575.1; -  
DR EMBL; K03226; AAC97138.1; -  
DR EMBL; K02286; AAA61252.1; -  
DR EMBL; A21571; CAA01559.1; -  
DR EMBL; A18397; CAA01390.1; -  
DR PIR; A00931; UKHU.  
DR PDB; 1KDU; 31-OCT-93.

Query Match 99.9%; Score 2298; DB 1; Length 431;  
 Best Local Similarity 99.8%; Pred. No. 1.7e-174;  
 Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFNSIHCNCPKFGGQHCIDKSKTCYEGNGHYRG 60  
 DB 21 SNELHQPSPNCDCCLNGGTCVSNKYFNSIHCNCPKFGGQHCIDKSKTCYEGNGHYRG 80  
 QY 61 KASDTMTGRCLPNSATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
 DB 81 KASDTMTGRCLPNSATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140  
 QY 121 PLVQECMVHCDKPKSSPPEELKFCQGGKTLPRPKIIGBFTTINOPWFAAIYRRH 180  
 DB 141 PLVQECMVHCDKPKSSPPEELKFCQGGKTLPRPKIIGBFTTINOPWFAAIYRRH 200  
 QY 181 RGSVTVVCGSLSPSCWVLSATHCFIDYPKEDYIVVLGSRSLNSNTQGMKEVENLI 240  
 DB 201 RGSVTVVCGSLSPSCWVLSATHCFIDYPKEDYIVVLGSRSLNSNTQGMKEVENLI 260  
 QY 241 LHKDYSADTLAHNDIALALKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCBITGFK 300  
 DB 261 LHKDYSADTLAHNDIALALKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCBITGFK 320  
 QY 301 ENSTDYLYPQLKMTVVKLSHRECOQPHYGYSEVTTMLCAADPQWKTDSCQDSSGGL 360  
 DB 321 ENSTDYLYPQLKMTVVKLSHRECOQPHYGYSEVTTMLCAADPQWKTDSCQDSSGGL 380  
 QY 361 VCSLQGRMTLTGIVSGRGKALDKDPGYTVVSHFLPWIRSHTEKENGAL 411  
 DB 381 VCSLQGRMTLTGIVSGRGKALDKDPGYTVVSHFLPWIRSHTEKENGAL 431

RESULT 2  
 UROK\_PAPCY STANDARD; PRT; 433 AA.

AC P16227;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
 DE (U-plasminogen activator).  
 GN P1AU.  
 OS Papio cynocephalus (Yellow baboon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Papio.  
 OX NCBI\_TaxID=9556;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=thoracic aorta;  
 RX MEDLINE=90287734; PubMed=2113276;  
 RA Au Y.P.T.; Wang T.W.; Clowes A.W.;  
 RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type  
 plasminogen activator";  
 RL Nucleic Acids Res. 18:3411-3411(1990).  
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 plasminogen to form plasmin.  
 CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS  
 OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A  
 LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW  
 MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: Contains 1 kringle domain.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 -----  
 CC EMBL; X51935; CAA36200.1; -.  
 CC PIR; S14687; UKBAY.  
 DR HSP; P00749; ILMW.  
 DR MEROPS; S01.231; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR006209; EGF like.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Ser. protease\_Try.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; FALSE NEG.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 KW Kringle; EGF-like domain; Zymogen; Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
 FT CHAIN 21 176 CHAIN A (BY SIMILARITY).  
 FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).  
 FT CHAIN 178 433 CHAIN B (BY SIMILARITY).  
 FT DOMAIN 26 62 EGF-LIKE.  
 FT DOMAIN 69 150 KRINGLE.  
 FT DOMAIN 151 177 CONNECTING PEPTIDE.  
 FT DOMAIN 178 433 SERINE PROTEASE.  
 FT DISULFID 30 38 BY SIMILARITY.  
 FT DISULFID 32 50 BY SIMILARITY.  
 FT DISULFID 52 61 BY SIMILARITY.  
 FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 208 224 BY SIMILARITY.  
 FT DISULFID 216 287 BY SIMILARITY.  
 FT DISULFID 315 384 BY SIMILARITY.  
 FT DISULFID 347 363 BY SIMILARITY.  
 FT ACT\_SITE 223 223 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 274 274 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 378 378 CHARGE RELAY SYSTEM.  
 FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).  
 SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match 93.4%; Score 2148; DB 1; Length 433;  
 Best Local Similarity 92.5%; Pred. No. 1.2e-162;  
 Matches 389; Conservative 17; Mismatches 10; Indels 4; Gaps 2;

QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFNSIHCNCPKFGGQHCIDKSKTCYEGNGHYRG 60  
 DB 21 SREL-QVPSDCGCLNGGTCVSNKYFNSIHCNCPKFGGQHCIDKSKTCYEGNGHYRG 79  
 QY 61 KASDTMTGRCLPNSATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
 DB 80 KASDTMTGRCLPNSATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 139  
 QY 121 PLVQECMVHCDKPKSSPPEELKFCQGGKTLPRPKIIGBFTTINOPWFAAIYRRH 180  
 DB 140 QVQECMVHCDKPKSSPPEELKFCQGGKTLPRPKIIGBFTTINOPWFAAIYRRH 199  
 QY 181 RGSVTVVCGSLSPSCWVLSATHCFIDYPKEDYIVVLGSRSLNSNTQGMKEVENLI 240  
 DB 200 RGSVTVVCGSLSPSCWVLSATHCFIDYPKEDYIVVLGSRSLNSNTQGMKEVENLI 259  
 QY 241 LHKDYSADTLAHNDIALALKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCBITG 297

```
Db 260 LHEDYSADTLAHNDIALKRSREGRCAQPSRTIQTICLPSMYNDPDPFGTSCBITG 319
Qy 298 FGKSTNDYLYPEQLKMTWVKLVSHRECOQPHYGVSEVTTKMLCAADPOWKTDSOQDSG 357
Db 320 FGKSTNDYLYPEQLKMTWVKLVSHKQCOQPHYGVSEVTTKMLCAADPOWKTDSOQDSG 379
Qy 358 GPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGAL 411
Db 380 GPLVCSIQGHMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGAL 433

RESULT 3
UROK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y.; Pearson D.; Altus M.S.; Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01648; CAA25806.1; -
DR EMBL; X02724; CAA26511.1; -
DR PIR; A00932; UKPG.
DR HSSP; P00749; IKDU.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS01186; EGF 2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
```

```
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 442
FT CHAIN 21 188
FT CHAIN 190 442
FT DOMAIN 29 65
FT DOMAIN 72 153
FT DOMAIN 154 189
FT DOMAIN 190 442
FT CARBOHYD 152 152
FT DISULFID 33 41
FT DISULFID 35 53
FT DISULFID 55 64
FT DISULFID 179 310
FT DISULFID 220 236
FT DISULFID 228 299
FT DISULFID 324 393
FT DISULFID 356 372
FT DISULFID 383 411
FT ACT_SITE 235 235
FT ACT_SITE 286 286
FT ACT_SITE 287 287
FT ACT_SITE 241 241
FT CONFLICT 242 242
FT CONFLICT 248 248
FT CONFLICT 288 288
SQ SEQUENCE 442 AA; 49116 MW; EB32FCEP501321EE CRC64;

Query Match 80.8%; Score 1859.5; DB 1; Length 442;
Best Local Similarity 79.3%; Pred. No. 8.3e-140;
Matches 334; Conservative 33; Mismatches 43; Indels 11; Gaps 2;

Qy 1 SHELQV--PSNCDCLNGGTCVSNKYFSNIHWCNPKKFGQGHCEIDSKSKCYEGNGHFPY 58
Db 21 SHELQBSGASNCGLNGGKCVSYKFSNIQRCSCPKKFQGEHCEIDTSQTCFENGHSY 80
Qy 59 RGKASTDTMGRCPLPWN SATVLQOITYHAHRS DALQLGLGKHN YCRNPNRRRPWCYVQVG 118
Db 81 RGKANTNTGRCPLPWN SATVLLNTYHAHRP DALQLGLGKHN YCRNPNRRRPWCYVQVG 140
Qy 119 LKPLVQECWHDCA-----DGKFPSSPEELKFCQCKOTLRPRPKITGGETTIE 169
Db 141 LKQLVQECWVNGSGSHR PAYDGNKPFSTPEKVFQCGKALRPRPKIVGGKSTTIE 200
Qy 170 QPWFAAIYRRHRGSGSVTVVCGSLISPCWVTSATHCFIDYPKKEDYIVLGRSLNSNTQ 229
Db 201 QPWFAAIYRRHRGSGSVTVVCGSLISPCWVVSATHCFINYQKEDYIVLGRQLHSSTH 260
Qy 230 GEMKFEVENLILHKDYADTLAHNDIALKIRSKRGCAQPSRTIQTICLPSMYNDPQF 289
Db 261 GEMKFEVEKLILHEDYSADSLAHNDIALKIRTDKGCAQPSRSIQTCILPPVNGDAHF 320
Qy 290 GTSCEITGFGKENSTDYLYPEQLKMTWVKLVSHRECOQPHYGVSEVTTKMLCAADPOWKT 349
Db 321 GASCEIVGFKEDPSDYLYPEQLKMTWVKLVSHRECOQPHYGVSEVTTKMLCAADPOWKT 380
Qy 350 DSCQDGGGLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENG 409
Db 381 DSCQDGGGLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENG 440
Qy 410 A 410
Db 441 A 441

RESULT 4
UROK_BOVIN STANDARD; PRT; 433 AA.
AC Q05589; Q28209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
```

FT	DISULFID	35	53	BY SIMILARITY.
FT	DISULFID	55	64	BY SIMILARITY.
FT	DISULFID	170	301	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	211	227	BY SIMILARITY.
FT	DISULFID	219	290	BY SIMILARITY.
FT	DISULFID	315	384	BY SIMILARITY.
FT	DISULFID	347	363	BY SIMILARITY.
FT	DISULFID	374	402	BY SIMILARITY.
FT	ACT_SITE	226	226	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	277	277	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	378	378	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	CONFLICT	189	189	A -> T (IN REF. 27).
SEQ	SEQUENCE	433 AA;	48730 MW;	4DE1B8DA4A7027A CRC64;

  

Query Match		76.8%;	Score 1767;	DB 1;	Length 433;
Best Local Similarity		75.1%;	Pred. No. 1.7e-132;		
Matches	310;	Conservative	45;	Mismatches	56;
				Indels	2;
				Gaps	1;
Qy	1	SNELHQV--PSNCDCNLNGGTCVSNKQSPPEELKFCQGGKTLRPFKIIGGEFTTINQPFWEAIIYR	178		
Db	21	SNEVHKESGBSNCGCLNGGKCVTKYFNSIQRCSCPKKFGQEHCEIDTSKTCYQGNHSHY	80		
Qy	59	RGKASDTMGRPCLPWNSATVLQOTTHAIRSDALQILGLKHNYCRNPDPNRRPWCYVQVG	118		
Db	81	RGKANRDLGRPCLAWDSPTVLLKMYAHARSDAQLGLGKHNYCRNPDPNQRPPWCYQIG	140		
Qy	119	LKPLVQECMVHDCADGKKSSPPPEELKFCQGGKTLRPFKIIGGEFTTINQPFWEAIIYR	178		
Db	141	LKQFVQFCMVQDCSVGKSPSPREKEEFQCGQKALRPFAIVGGQVTNAENQPFWEAIIYR	200		
Qy	179	RHRGGSVTVYCGGLSLISPCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKPFVEN	238		
Db	201	RHRGGSVTVYCGGLSLISPCWVVSATHCFIDHPKKENIVYVYLGSRSLNSNDRGEMQFEVK	260		
Qy	239	LIILHKYSADTLAHHNDIALLLKIRSEKRCQAQPSRTIQICLPSMWNDFQFGTSCEITGF	298		
Db	261	LIILHEDYSAEFLAHHNDIALLLKIRTSRGCAQPSRSIQICLPPHEHDAHSRTRCEITGF	320		
Qy	299	GKENSITDYLPEOLKMTVVKLIISHRCQOPHYVYGVSEVTTKMLCAAPQWKTDCSQGSGG	358		
Db	321	GKENSIDRYSDLEKMTFVLSVSHVEVQOPHYVYGAETDKMLCAAPQWETDCSQGSGG	380		
Qy	359	PVCSLQGRMTLTGTVSWGRGCALKDQKPGVYTVRSVHFLPWIRSHTKBENGAL	411		
Db	381	PLVCTIQGRLLTGTIVSGWEDCAMKYKPGVYTVRSVHFLPWINTHTRGEINLV	433		

  

RESULT 5	UROC RAT	STANDARD;	PRT;	432 AA.
AC	P29538;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)			
DE	(U-plasminogen activator).			
GN	PLAU.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
NCBI	TaxId=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Fischer 344;			
RX	MEDLINE=92233409; PubMed=1568219;			
RA	Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,			
RA	Keifford R.F.;			
RT	"Transcriptional and posttranscriptional activation of urokinase			
RT	plasminogen activator gene expression in metastatic tumor cells.";			
RL	Cancer Res. 52:2489-2496(1992).			
RN	[2]			
RC	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			



RA Rabbani S.A.; to the EMBL/GenBank/DBJ databases.  
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH  
 CC CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM  
 CC CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE  
 CC LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X63434; CAA45028.1; -;  
 DR EMBL; X63651; CAA46601.1; -;  
 DR PIR; S24604; S18932.  
 DR HSP; P00749; 1KDU.  
 DR MEROPS; S01\_231; -;  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR006210; EGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Ser. Protease\_Try.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 KW Kringle; EGF-like domain; Zymogen; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
 FT CHAIN 20 177 CHAIN A (BY SIMILARITY).  
 FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).  
 FT CHAIN 179 432 CHAIN B (BY SIMILARITY).  
 FT DOMAIN 27 63 EGF-LIKE.  
 FT DOMAIN 70 151 KRINGLE.  
 FT DOMAIN 152 178 CONNECTING PEPTIDE.  
 FT DOMAIN 179 432 SERINE PROTEASE.  
 FT DOMAIN 31 39 BY SIMILARITY.  
 FT DISULFID 33 51 BY SIMILARITY.  
 FT DISULFID 53 62 BY SIMILARITY.  
 FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 210 226 BY SIMILARITY.  
 FT DISULFID 218 289 BY SIMILARITY.  
 FT DISULFID 314 383 BY SIMILARITY.  
 FT DISULFID 346 362 BY SIMILARITY.  
 FT DISULFID 373 401 BY SIMILARITY.  
 FT ACT\_SITE 225 225 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 276 276 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 377 377 CHARGE RELAY SYSTEM.  
 FT CONFLICT 16 16 N -> H (IN REF. 2).  
 FT CONFLICT 24 24 E -> G (IN REF. 2).  
 FT CONFLICT 332 332 D -> N (IN REF. 2).  
 SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match 73.5%; Score 1690.5; DB 1; Length 432;  
 Best Local Similarity 73.2%; Pred. No. 1.9e-126;  
 Matches 295; Conservative 47; Mismatches 60; Indels 1; Gaps 1;  
 Cyt 9 SNCDCLNGTCVSNKYESNTHWCNCPKFKGOHCEIDKSKTCYEGNHFYRGKASTDTMG 68  
 Db 29 SNOCQNGGVCSYKYFSSIRRCSCPKFKGEHCEIDTSKTCYHNGQSVYRGKANTDTKG 88  
 Qy 69 RPCLPWNASVLOQTYHAHSDALQLGLGHKNYCRNPNRRRPPWCYQVGLKPLVQECMV 128  
 Db 89 RPCLAWNPAVLOQTYNAHSDALSGLGHKNYCRNPNQRRPWCYVQIGLKQFVQECMV 148  
 Qy 129 HDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPMFAAIYRHRGGS-VTY 187  
 Db 149 QDCSLSKKPSSTVDQOQFCQKALRPRFKIVGGEFTWENQPMFAAIYLNKNGKSPSP 208  
 Qy 188 VCGSLSPCWVTSATHCFTDYPKKEDYIVYLGESRLNSNTQGMKEFEVENLILHKDYSA 247  
 Ph 209 KCGSLSPCWASATHCFVNPQKKEVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 268  
 Qy 248 DTLAHNDIALKIRSKEGRCAPSRITQITICLPSMYNDPQFGTSCETIGFGENSTDYL 307  
 Db 269 ETIAFHNDIALLKIRSTGCAQPSRTIQTICLPFGDAPFGSDCHITFGQESATDYF 328  
 Qy 308 YPQLQMTVVVKLISHRECQPHYGVSEVTTKMLCAADPQWKTSCQGDSCGSLVCSLQGR 367  
 Db 329 YPKDKMSVVVKIISHQCKQPHYGVSEINVKMLCAADPEWKTSCSGDSCGSLICNIDGR 388  
 Qy 368 MTLTGIVSWGCGCALDKPGVTRVSHFLPWISHTKEENGLA 410  
 Db 389 PTLGIVSWGCGCAERKNPGVTRVSYFLNWIQSHIGEENGLA 431  
 RESULT 6  
 UROK\_MOUSE  
 ID UROK\_MOUSE STANDARD; PRT; 433 AA.  
 AC P06869;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
 DE (U-plasminogen activator).  
 GN PLAU.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85179474; PubMed=2985383;  
 RA Belin D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y.,  
 RA Reich E., Kocher H.P., Duvoisin R.M.;  
 RT "Cloning, nucleotide sequencing and expression of cDNAs encoding  
 RT mouse urokinase-type plasminogen activator.";  
 RL Eur. J. Biochem. 148:225-232 (1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88163489; PubMed=2831940;  
 RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;  
 RT "The murine urokinase-type plasminogen activator gene.";  
 RL Biochemistry 26:8270-8279 (1987).  
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS  
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A  
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE LOW  
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -



```

DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 171 CHAIN A (BY SIMILARITY).
FT CHAIN 173 434 CHAIN B (BY SIMILARITY).
FT DOMAIN 36 72 EGF-LIKE.
FT DOMAIN 79 158 KRINGLE.
FT DOMAIN 159 172 CONNECTING PEPTIDE.
FT DOMAIN 173 434 SERINE PROTEASE.
FT DISULFID 40 48 BY SIMILARITY.
FT DISULFID 42 60 BY SIMILARITY.
FT DISULFID 62 71 BY SIMILARITY.
FT DISULFID 162 296 INTERCHAIN (BY SIMILARITY).
FT DISULFID 202 218 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 210 285 BY SIMILARITY.
FT DISULFID 310 379 BY SIMILARITY.
FT DISULFID 342 358 BY SIMILARITY.
FT DISULFID 369 397 BY SIMILARITY.
FT ACT SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 373 373 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 228 228 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 434 AA; 49400 MW; BD881048DD666A55 CRC64;

Query Match 43.6%; Score 1004; DB 1; Length 434;
Best Local Similarity 47.4%; Pred. No. 3.7e-72;
Matches 188; Conservative 64; Mismatches 123; Indels 22; Gaps 6;

QY 11 CDCLEGGTCVSNKFSNHWNCNPKFGGCGCEIDKSKTCYEGNGHFYRGKASDTWGRP 70
DB 40 CQCLNGGTCTIYRFSQIKRCLCEGYGLHCEIDTNSICYSNGEDYRGWAEDP----G 95
QY 71 CLPNSATVLO-QTYHAHRSDALQLGLGKHNCRPNDRPFCWVQVGLKPLVQCMVH 129
DB 96 CLYWDHPSVIRWDYHADLKNALQLGLGKHNCRPNDRPFCWVQVGLKPLVQCMVH 149
QY 130 CDADGKPPSPPELKCQGGKTLRPRKIIGFETIENPFAIYRHRGSGVTVVC 189
DB 150 -----TPCSTIERCETCGORSFSKYFKIVGGQAEVETQWIAQIPQIM-GTDQFLC 202
QY 190 GGSLLSPCWITSAFHCFTD---VPKEDYIVVLGRSLNNTQGMKFEVENILHKDY 245
DB 203 GGSLLDFCWLTAAHCFNPTKQPNKSVYKVFLOKSLINTNDEHQVFWFDEIISHPDF 262
QY 246 SADTLAHNHDIALKIRSKRCAQPSRTIQTICLPSMYNDPQFTGCEITFGKENSST 305
DB 263 TDHTGGNDNDIALIRIRATASQCAVESNYVTVCLPEKNLNLNDTWCETIAGYQKNSYD 322
QY 306 YLYPEQLKMTVVKLISHRECOOPHYGSEVTTKMLCAADPOWKTDCQGSQGLVCSLQ 365
DB 323 IYFAQLMSATVNLSIQDCKNKYITSTRVTDNVMVCAGDPLWETDACKGDSGGPMVCEHN 382
QY 366 GRMTLTGIVSWGRGALKDKPGVTVRYSHFLPWTIRSH 402
DB 383 GRMTLYGIVSWGDCCKAKNKGVTYTRYLNWIDSN 419

RESULT 8
ID _URT2_DESRO STANDARD; PRT; 477 AA.
AC P15638;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA
DE alpha-2) (BAT-PA) (T-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;

```

```

Desmodontinae; Desmodus.
OC NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Salivary gland;
RX MEDLINE=920390367; PubMed=1937019;
RA Kraetschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX TISSUE=Salivary gland;
RX MEDLINE=90036867; PubMed=2509450;
RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
RT "Isolation, characterization, and cDNA cloning of a vampire bat
RT salivary plasminogen activator.";
RL J. Biol. Chem. 264:17947-17952(1989).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RX Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC THE PRESENCE OF FIBRIN I.
CC -!- SUBUNIT: Monomer.
CC -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC STIMULATION OF ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63988; AAA31593.1; -.
DR EMBL; J05082; AAA31596.1; -.
DR PIR; J05098; J05098.
DR HGSP; P98119; IAS1.
DR MEROPS; S01.232; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.

```

```

DR SMART; SMO0130; KR; 1.
DR SMART; SMO0020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 477
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 225 477
FT ACT_SITE 272 321
FT ACT_SITE 321 321
FT ACT_SITE 428 428
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 214 345
FT DISULFID 257 273
FT DISULFID 265 334
FT DISULFID 359 434
FT DISULFID 391 407
FT DISULFID 424 452
FT CARBOHYD 185
FT CARBOHYD 398
FT CONFLICT 403 403
FT CONFLICT 417 417
FT CONFLICT 435 435
SQ SEQUENCE 477 AA; 53719 MW; 17486555C0E5077C CRC64;

Query Match 37.6%; Score 864.5; DB 1; Length 477;
Best Local Similarity 42.9%; Pred. No. 4.4e-61;
Matches 178; Conservative 60; Mismatches 148; Indels 29; Gaps 9;

QY 3 ELHQP-----SNCDGNGCTVSNKYFSNIHWCNCPKFGQCHCEIDKSKTCYEGNGHYF 58
DB 78 QCHTVPVKSCSELRCFNGCTCQAAASFSDF-VCQCPKGYTGQCEVDTHATCYKQDQGVY 136
QY 59 RGKASTDTMGRCLPNSATVLOQTYHAHRSDALQGLGKHNKCNPNRRPWCYQVVG 118
DB 137 RGTWSTSGAOCINNSNLLTRTYNGRRSDAITLGLGNHNYCNPNNSKPCWCYVKA 196
QY 119 LKPLVQECWHDCAQKGPSPPELKFQCG-QKTLRPRFKIIGGEFTTIENQWFAAIY 177
DB 197 SKFILEFCVPVCS-----KATCGLRKVKEPQLHSTGGLPTDITSHPWQAIF 244
QY 178 RRRHGS-VTVYCGSLISPCWVIAATHCFD-YPKEDYIVYGRSLNNTQGMKFE 235
DB 245 AQNRSSGGRFLCGGILITSSCWLVTAACHFOERYPQHLRVV-LGRTVRVRPGKEQTFE 303
QY 236 VENLILHKDYSADTLAHHNDIALIKRSKGRCAQPSRTIQTICLPSMYNDPQFQTSCEI 295
DB 304 VEKCVHEFFDDT--YNNDIALLQKSGSPCAQESVRAICLPEANILQLPDWTCEL 361
QY 296 TFGKENSTDYLYPQLKMTVVKVLSHRECOQPHYGVSEVTKMLCAADPOMKT----- 349
DB 362 SGYGHKSSPFYSEQLKEGHRVLPSSRCTSKFLFNKVTNNMLCAGDTRSGEYIPNVH 421
QY 350 DSCQDGGGLVCSLQGRMTLGIYSGRGKALCKPKGYTVFVSHFLWIRSHTK 404
DB 422 DACQDGGGLVCMNDNHTLGLIISWGVGCGEKDIPGYTVTKVTNYLGMWRDNR 476

```

## RESULT 9

```

TPA_HUMAN
ID TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retelplase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221(1983).
[2]
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saico Y., Hayashi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695(1988).
[3]
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
RA Hsiung N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors.";
RL DNA 6:461-472(1987).
[4]
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Friesner Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
RL J. Biol. Chem. 261:6972-6985(1986).
[5]
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
[6]
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,
RA Odenakker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
RL Mol. Biol. Med. 3:279-292(1986).
[7]
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Umbilical vein;
RX MEDLINE=30192129; PubMed=2107528;
RA Siebert P.D., Fong K.;
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
RT human endothelial cells.";
RL Nucleic Acids Res. 18:1086-1086(1990).
[8]
RN [8]

```





Db 209 FLOGGILISSCWLTAAHCFQERYPPQHLRV-LGRTYRVKPKBQTEFEVEKCIHEEF 267  
Qy 246 SADTLAHHNDIALKIRSKRCAQPSRTIQTICLPMSYNDPQGTSCITGFGKENSTD 305  
Db 268 DDDT--YNNDIALLQLKSPQCAQESDVAICLPEANQLPDTWTECLSGYKHKSS 325  
Qy 306 YLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWKT-----DSQCQSDGGP 359  
Db 326 PFYSEQLKEGHVLYPSSRCTSKFLFNKTVTNMMLCAGDTRSGEITYPNVHDAQCGSDGGP 385  
Qy 360 LVCSLQGRMTLTGIVSGRGKALKDXPGVTVRYSHFLPWIRSHTK 404  
Db 386 LVCNDNHMTLLGIISGVGCGEKDIEGVTKVNYLGLWIRDNR 430

RESULT 11  
TPA RAT  
ID ID TPA RAT STANDARD; PRT; 559 AA.  
AC P19637;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-NOV-1991 (Rel. 20, last sequence update)  
DT 28-FEB-2003 (Rel. 41, last annotation update)  
DE Tissue-type plasminogen activator precursor (RC 3.4.21.68) (tPA)  
DE (t-PA) (t-plasminogen activator).  
GN PLAT.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89170114; PubMed=3148445;  
RA Ny T.; Leonardson G.; Hsueh A.J.W.;  
RT Cloning and characterization of a cDNA for rat tissue-type  
RL plasminogen activator."  
RL DNA 7:671-677(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90130448; PubMed=2105315;  
RA Feng P.; Ohlsson M., T.;  
RT "The structure of the TPA-less rat tissue-type plasminogen activator  
RT gene. Species-specific sequence divergences in the promoter predict  
RT differences in regulation of gene expression."  
RL J. Biol. Chem. 265:2022-2027(1990).  
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN  
CC TO ELASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY  
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT  
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND  
CC MANY OTHER PHYSIOPATHOLOGICAL EVENTS.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE  
CC BOND.  
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.  
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER  
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER  
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.  
CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A  
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 2 kringle domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
EMBL; M23697; AAA41812.1; -.

DR EMBL; M31197; AAA42261.1; -.  
DR EMBL; M31185; AAA42261.1; JOINED.  
DR EMBL; M31186; AAA42261.1; JOINED.  
DR EMBL; M31187; AAA42261.1; JOINED.  
DR EMBL; M31188; AAA42261.1; JOINED.  
DR EMBL; M31189; AAA42261.1; JOINED.  
DR EMBL; M31190; AAA42261.1; JOINED.  
DR EMBL; M31191; AAA42261.1; JOINED.  
DR EMBL; M31192; AAA42261.1; JOINED.  
DR EMBL; M31193; AAA42261.1; JOINED.  
DR EMBL; M31194; AAA42261.1; JOINED.  
DR EMBL; M31195; AAA42261.1; JOINED.  
DR EMBL; M31196; AAA42261.1; JOINED.  
DR EMBL; A35029; A35029.  
DR HSSP; P00750; IRTF.  
DR MEROPS; S01.232; -.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000083; Fibronctnl.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF00039; fnl; 1.  
DR Pfam; PF00051; kringle; 2.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 2.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00058; FN1; 1.  
DR SMART; SM00130; KR; 2.  
DR SMART; SM00020; Tryp\_spc; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 2.  
DR PROSITE; PS00070; KRINGLE\_2; 2.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SRR; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.  
FT SIGNAL 1 17 PROBABLE.  
FT PROPEP 18 29  
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR A  
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A  
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B  
FT CHAIN 36 78 FIBRONECTIN TYPE-I.  
FT DOMAIN 79 117 EGF-LIKE.  
FT DOMAIN 124 205 KRINGLE 1.  
FT DOMAIN 213 294 KRINGLE 2.  
FT DOMAIN 309 559 SERINE PROTEASE.  
FT ACT\_SITE 355 355 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 404 404 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 510 510 CHARGE RELAY SYSTEM.  
FT DISULFID 38 68 BY SIMILARITY.  
FT DISULFID 65 75 BY SIMILARITY.  
FT DISULFID 83 94 BY SIMILARITY.  
FT DISULFID 88 105 BY SIMILARITY.  
FT DISULFID 107 116 BY SIMILARITY.  
FT DISULFID 124 205 BY SIMILARITY.  
FT DISULFID 145 187 BY SIMILARITY.  
FT DISULFID 176 200 BY SIMILARITY.  
FT DISULFID 213 234 BY SIMILARITY.  
FT DISULFID 234 276 BY SIMILARITY.  
FT DISULFID 265 289 BY SIMILARITY.  
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 340 356 BY SIMILARITY.  
FT DISULFID 348 417 BY SIMILARITY.







RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences.";  
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN  
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY  
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT  
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND  
 CC MANY OTHER PHYSIOLOGICAL EVENTS.  
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE  
 CC BOND.  
 CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.  
 CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER  
 CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER  
 CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.  
 CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A  
 CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -!- SIMILARITY: Contains 2 kringle domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC ENBL; J03520; AAA0470.1; --  
 CC ENBL; BC011256; AAH1256.1; --  
 CC PLR; A29941; A29941.  
 CC HSP; P00750; IAS5.  
 CC MEROPS; S01.232; --  
 CC MGD; MGI:97610; Plat.  
 CC InterPro; IPR001314; Chymotrypsin.  
 CC InterPro; IPR006209; EGF\_like.  
 CC InterPro; IPR000083; Fibronctnl.  
 CC InterPro; IPR006210; IEGF.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR001254; Ser\_protease\_Try.  
 CC Pfam; PF00008; EGF; 1.  
 CC Pfam; PF00039; fnl; 1.  
 CC Pfam; PF00051; kringle; 2.  
 CC Pfam; PF00089; trypsin; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC PRINTS; PR00018; KRINGLE.  
 CC ProDom; PD000395; Kringle; 2.  
 CC SMART; SM00181; EGF; 1.  
 CC SMART; SM00058; fnl; 1.  
 CC SMART; SM00130; KR; 2.  
 CC SMART; SM00020; Tryp\_Spc; 1.  
 CC PROSITE; PS00022; EGF 1; 1.  
 CC PROSITE; PS01186; EGF 2; 1.  
 CC PROSITE; PS01253; FIBRONECTIN 1; 1.  
 CC PROSITE; PS00021; KRINGLE 1; 2.  
 CC PROSITE; PS00070; KRINGLE 2; 2.  
 CC PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC plasminogen activation: Hydrolase; Serine protease; Glycoprotein;  
 KW Plasma; Kringle; EGF-like domain; Repeat; Signal.  
 FT SIGNAL 1 17  
 FT PROPEP 18 29 PROBABLE.  
 FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.  
 FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A  
 FT CHAIN 309 559 CHAIN.  
 FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B  
 FT CHAIN 309 559 CHAIN.  
 FT DOMAIN 36 78 FIBRONECTIN TYPE-I.

FT	DOMAIN	79	117	EGF-LIKE.
FT	DOMAIN	124	205	KRINGLE 1.
FT	DOMAIN	213	294	KRINGLE 2.
FT	DOMAIN	309	559	SERINE PROTEASE.
FT	ACT_SITE	355	355	CHARGE RELAY SYSTEM.
FT	ACT_SITE	404	404	CHARGE RELAY SYSTEM.
FT	ACT_SITE	510	510	CHARGE RELAY SYSTEM.
FT	DISULFID	38	68	BY SIMILARITY.
FT	DISULFID	66	75	BY SIMILARITY.
FT	DISULFID	83	94	BY SIMILARITY.
FT	DISULFID	88	105	BY SIMILARITY.
FT	DISULFID	107	116	BY SIMILARITY.
FT	DISULFID	124	205	BY SIMILARITY.
FT	DISULFID	145	187	BY SIMILARITY.
FT	DISULFID	176	200	BY SIMILARITY.
FT	DISULFID	213	294	BY SIMILARITY.
FT	DISULFID	234	276	BY SIMILARITY.
FT	DISULFID	265	289	BY SIMILARITY.
FT	DISULFID	297	428	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	340	356	BY SIMILARITY.
FT	DISULFID	348	417	BY SIMILARITY.
FT	DISULFID	442	516	BY SIMILARITY.
FT	DISULFID	474	490	BY SIMILARITY.
FT	DISULFID	506	534	BY SIMILARITY.
FT	CARBOHYD	149	149	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	481	481	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	260	260	G -> A (IN REF. 1).
FT	CONFLICT	325	325	P -> A (IN REF. 1).
SQ	SEQUENCE	559 AA;	63122 MW;	8CCE2BDB94514D9 CRC64;

Query Match 35.9%; Score 826.5; DB 1; Length 559;  
 Best Local Similarity 35.7%; Pred. No. 5.3e-58;  
 Matches 176; Conservative 65; Mismatches 155; Indels 97; Gaps 10;

QY	3	ELHOVP----	SNCDCLNGTFCVSNKYFSNTHWNCPCFKGSGHCEIDKSKTCYEGNGHFY	58
Db	74	QCHSVPRVSCSEPRCNGTCCQALFSDP	VCQPDGFVGKRCIDITRATCFEEQITY	132
QY	59	RGAKASTDTMGRPCLPWNSATVLOOTYAHASDALQLGLGKHNYCRNPNRRRPMCVQVG	118	
Db	133	RGWTSTAESGAECINWNSSVLSLKPYNARRPNAIKLGLGNHNYCRNPDRLDKPWCYFKA	192	
QY	119	LKPLVECMVHDCADGKPS-----	138	
Db	193	GKYTEFCSTPACPKGKSEDCYVGKGYVYRGTHSLTTSQASCLPWNSIVLMGKSYTAWRT	252	
QY	139	-----SPEELK-----	162	
Db	253	NSQALGLGRHNYCRNPDGDARPMCHWKORKLTWEYCDMSPCSTCGURQYKRPQFRKGG	312	
QY	163	EFTTIENQPMFAAIY-RRHRGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDYIVYVLR	221	
Db	313	LYTDITSHWPQAPIFVKNKRSPGERFLCGGLISSCWLSAAHCFLEFPFNHLKVVLR	372	
QY	222	SRLSNTOGEMKEVENLILHKDYSADTLAHNDIALKLRSEKGCQAQPSRTIOTICLP	281	
Db	373	TYRVFGEBEQTEIEIKYIVHEEFDDET--YNDNIALQLRSQSKQCAQESSVGTACLP	430	
QY	282	SMYNDPQF-----GTSCEITGFGKENSTDYLYPBQLKMTVVVKLISHRECQPHYVGSVTT	337	
Db	431	-----DENLQLPDWTECELSGYGKHEASSFFSDRLKEAHVRLYPSRSCRTSQHLFNKTVTN	486	
QY	338	KMLCAADP-----QWKTDSCQGDGSGPLVCSLQGRMTLTGIYVSWRGCAKDKPQGVYTRV	392	
Db	487	NMLCAGDTRSGGNQDLHDACQGDGSGPLVCMINKMTLTLGIISWGLGCGQKDPGVYTKV	546	
QY	393	SHPLPWIRSHTKE	405	
Db	547	TNYLDWIHDNMKQ	559	

RESULT 14  
 TPA\_BOVIN

```
ID TPA_BOVIN STANDARD; PRT; 566 AA.
AC Q28198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
and tPA.";
RL Int. Dairy J. 5:605-617 (1995).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZIMOGEN PLASMINOGEN
TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-314 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 2 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR HSSP; P00750; IRTF.
DR MEROPS; S01.232; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF-2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 2.
```

```
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT PROPEP 22 33 BY SIMILARITY.
FT CHAIN 34 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 34 314 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
CHAIN.
FT CHAIN 315 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
CHAIN.
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.
FT DOMAIN 83 121 EGF-LIKE.
FT DOMAIN 128 209 KRINGLE 1.
FT DOMAIN 219 300 KRINGLE 2.
FT DOMAIN 315 566 SERINE PROTEASE.
FT ACT_SITE 361 361 CHARGE RELAY SYSTEM.
FT ACT_SITE 410 410 CHARGE RELAY SYSTEM.
FT ACT_SITE 517 517 CHARGE RELAY SYSTEM.
FT DISULFID 42 72 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
FT DISULFID 149 191 BY SIMILARITY.
FT DISULFID 180 204 BY SIMILARITY.
FT DISULFID 219 300 BY SIMILARITY.
FT DISULFID 240 282 BY SIMILARITY.
FT DISULFID 271 295 BY SIMILARITY.
FT DISULFID 303 434 INTERCHAIN (BY SIMILARITY).
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 354 423 BY SIMILARITY.
FT DISULFID 448 523 BY SIMILARITY.
FT DISULFID 480 496 BY SIMILARITY.
FT DISULFID 513 541 BY SIMILARITY.
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BE4E32276C3 CRC64;

Query Match 35.5%; Score 818; DB 1; Length 566;
Best Local Similarity 36.3%; Pred. No. 2.5e-57;
Matches 179; Conservative 62; Mismatches 158; Indels 92; Gaps 11;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFVSNHWNCNPKKFGQCEIDKSKTCYEGNGHY 58
DB 78 QCHSVPRVSCSEPCWFCNGGTCRQALYSSDF-VQCPEGFNGKLCEDATATCYKQGVAY 136

QY 59 RGKASTDTMGRPCLPWN SATVLQOTYAHRS DALQLGLGHKHN YCRNPDNR RRRCYVQVG 118
DB 137 RGTWSTAESGAECANWNSGLAMKPYSGRRPNAIRLGLGNHNYCRNPDQDSKEWYVKA 196

QY 119 LKPLVQBCMVHDC-----DGK----- 135
DB 197 GKVISFCSTPACAKVAEDGDGCTGNGLA YRTRSH TSGASCLPWN SVFLT SKIYTAW 256

QY 136 KPSSP-----PEELKFQ-----CG-QKTLRPRFKII 160
DB 257 KSNAPALGLGKHNCRNPDQAQPCHVWKDRQLTW EYCDVPCVTCGLRQYKFPQRIK 316

QY 161 GGEFTTIENQPF AAIY-RRHGGSVYTVCGGSLISPCWVISATHCFIDYPKKEDYIVYL 219
DB 317 GGLFADITSHPMQAAIFVKNRRSPGERFLCGGILISSCWVLAAAHCFQERYPPHLLKVFL 376

QY 220 GRSRLNSNTQGEKMFVENLILHKYSA DTLAHNDIAL LKIRSK EGRCAQSRITQ TIC 279
DB 377 GRTYRLVPGEEQTFEVEKYLHKFDDDT--YNDIAL LHLKSLD TCAESASVRTIC 434

QY 280 LPSMYNDPQGTSCETGFGKENSTDYLYPEQLKMTVVKLISHRECCQPHYTGSEVTTKM 339
DB 435 LPDASLQLPDWTECELSGYGKHSSPFPFSERLKEAHVRLYPSRCTSQHLFNRTVTNNM 494
```

```

QY 340 LCAADPOM-----KTDSCQSGSGGLVCSLQGRMTLTGIVSWGRGCALKKPKGVYTRVS 393
DB 495 LCAGTRGGGHTNLHDACQCGSGGLVCMKDNHMLVGLISWLGCGCKKQVPGVYTKVT 554
QY 394 HPLPWIRSHK 404
DB 555 NYLDWIRDNR 565

RESULT 15
ID URTG DESRO STANDARD; PRT; 394 AA.
AC P49150;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA
DE gamma).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RL vampire bat Desmodus rotundus: cloning and expression.";
RN Gene 105:229-237(1991).
RP [2]
RP CHARACTERIZATION.
RA MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M63990; AAA31595.1; -.
CC PIR: J50600; JS0600.
CC HSP: P98119; IA51.
CC MEROPS: S01.239; -.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR001254; Ser. protease_Try.
CC Pfam: PF00051; kringle_1.
CC Pfam: PF00089; trypsin; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC PRINTS: PR00018; KRINGLE.
CC ProDom: PD000395; Kringle; 1.
CC SMART: SM00130; KR; 1 Spc; 1.
CC SMART: SM00020; Tryp Spc; 1.
CC PROSITE: PS00021; KRINGLE_1; 1.

```

```

DR PROSITE; PSS0070; KRINGLE_2; 1.
DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; Signal; Multigene family;
FT SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT DOMAIN 45 126 KRINGLE.
FT DOMAIN 142 394 SERINE PROTEASE.
FT ACT_SITE 189 189 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 45 126 BY SIMILARITY.
FT DISULFID 66 108 BY SIMILARITY.
FT DISULFID 97 121 BY SIMILARITY.
FT DISULFID 131 262 BY SIMILARITY.
FT DISULFID 174 190 BY SIMILARITY.
FT DISULFID 182 251 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.
FT DISULFID 308 324 BY SIMILARITY.
FT DISULFID 341 369 BY SIMILARITY.
FT CARBOHYD 315 315 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 394 AA; 44105 MW; 9CCD6F52F3D81FCD CRC64;

Query Match 32.7%; Score 752; DB 1; Length 394;
Best Local Similarity 42.0%; Pred. No. 2.8e-52;
Matches 155; Conservative 56; Mismatches 134; Indels 24; Gaps 7;

QY 45 DKSITTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRN 104
DB 40 DPHATCYKDOQVYRGTWSTSESGAQCINMNSNLLIRTYNGRMPEAVKGLGNHNYCRN 99
QY 105 PDNRPRWCYVQVGLKPLVQCMVHDCADGKPKSPPEELKFCQG-KQTLRPRFKIIGGE 163
DB 100 PDGASKWCYVIKARKTSESVCSPVCS-----KATCLRYKEPQLHSTGGL 147
QY 164 FTTTENOPWFAAIYRRHRGGS-VTVYCGGSLISPCWISATHCFID-YPKEDYIVYLR 221
DB 148 FTDITSHPWQAAIFAQNRRSSEGERFLOGGILISSCVLTAACHCFQERYPPQHLRV-LGR 206
QY 222 SRLNSNTGEMKFEVENILHDKYSADTLAHNDIALKIRSKGRCQAPRTQTICLP 281
DB 207 TYRVKPKGEQTFEVEKIVHEEFDDDT--YNDIALQLKSGSPQCAQESDVRALICLP 264
QY 282 SWYNDPQGTSCETGFGKENSTDYLPQOLKMTWKLIHRECOQPHYGVSEVTKMLC 341
DB 265 EANLQLPDWTECELSGYGKHKSSSPFYSEQLKEGHVRLYPSRCTSKFLFKNTVNNMLC 324
QY 342 AADPOWKT-----DSQSGSGGLVCSLQGRMTLTGIVSWGRGCALKKPKGVYTRVSHF 395
DB 325 AGDTRSGEIVPNVHDACQSGGLVCMNDNHTLLGIISWVGCGEKDIPGVYTKVTNY 384
QY 396 LPWIRSHK 404
DB 385 LGWIRDNR 393

RESULT 16
HGFA HUMAN
ID HGFA HUMAN STANDARD; PRT; 655 AA.
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Liver, and Serum;  
RX MEDLINE=93252878; PubMed=7693665;  
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,  
RA Kitamura N.;  
RT "Molecular cloning and sequence analysis of the cDNA for a human  
RT serine protease responsible for activation of hepatocyte growth  
RT factor. Structural similarity of the protease precursor to blood  
RT coagulation factor XII.";  
RL J. Biol. Chem. 268:10024-10028(1993).  
RN [2]  
RP SEQUENCE OF 40-655 FROM N.A.  
RA Zhao S., Odell C.;  
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY  
CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.  
CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A  
CC DISULFIDE BOND.  
CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN  
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.  
CC -!- TISSUE SPECIFICITY: LIVER.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: Contains 2 EGF-like domains.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D14012; BAA03113.1; -;  
CC EMBL; Z69923; CAA93803.1; -;  
CC PIR; A46688; A46688.  
CC HSP; P00763; IDPO.  
CC MEROPS; S01.228; -;  
CC Genew; HGNC:4894; HGFAC.  
CC MIM; 604552; -;  
CC GO; GO:0005576; C:extracellular; TAS.  
CC GO; GO:0006508; P:proteolysis and peptidolysis; TAS.  
CC InterPro; IPR001314; Chymotrypsin.  
CC InterPro; IPR00742; EGF\_2.  
CC InterPro; IPR006209; EGF-like.  
CC InterPro; IPR000083; Fibrinctn1.  
CC InterPro; IPR000562; FN Type II.  
CC InterPro; IPR006210; IEGF.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR001254; Ser\_protease\_Try.  
CC Pfam; PF00008; EGF; 2.  
CC Pfam; PF00039; fn1; 1.  
CC Pfam; PF00040; fn2; 1.  
CC Pfam; PF00051; kringle; 1.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PF00722; CHYMOTRYPSIN.  
CC PRINTS; PF00013; FNYPEII.  
CC PRINTS; PF00018; KRINGLE.  
CC ProDom; PD000995; FN Type II; 1.  
CC ProDom; PD000395; Kringle; 1.  
CC SMART; SM00181; EGF; 2.  
CC SMART; SM00058; FN1; 1.  
CC SMART; SM00059; FN2; 1.  
CC SMART; SM00130; KR; 1.  
CC SMART; SM00020; Tryp SPc; 1.  
CC PROSITE; PS00022; EGF\_1; 2.  
CC PROSITE; PS01186; EGF\_2; 1.  
CC PROSITE; PS01253; FIBRONECTIN\_1; 1.  
CC PROSITE; PS00023; FIBRONECTIN\_2; 1.  
CC PROSITE; PS00021; KRINGLE\_1; 1.

DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;  
FT EGF-like domain; Repeat; Zymogen.  
FT SIGNAL 1 30  
FT PROPEP 31 372 CLEAVED IN ACTIVE FORM.  
FT CHAIN 373 407 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT  
FT CHAIN 408 655 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG  
FT CHAIN 655 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG  
FT DOMAIN 108 FIBRONECTIN TYPE-II.  
FT DOMAIN 160 EGF-LIKE 1.  
FT DOMAIN 200 FIBRONECTIN TYPE-I.  
FT DOMAIN 241 EGF-LIKE 2.  
FT DOMAIN 286 KRINGLE.  
FT DOMAIN 408 SERINE PROTEASE.  
FT ACT\_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 598 598 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 108 133 BY SIMILARITY.  
FT DISULFID 122 148 BY SIMILARITY.  
FT DISULFID 164 175 BY SIMILARITY.  
FT DISULFID 169 186 BY SIMILARITY.  
FT DISULFID 188 197 BY SIMILARITY.  
FT DISULFID 202 230 BY SIMILARITY.  
FT DISULFID 228 237 BY SIMILARITY.  
FT DISULFID 245 256 BY SIMILARITY.  
FT DISULFID 250 267 BY SIMILARITY.  
FT DISULFID 269 278 BY SIMILARITY.  
FT DISULFID 286 367 BY SIMILARITY.  
FT DISULFID 307 349 BY SIMILARITY.  
FT DISULFID 338 362 BY SIMILARITY.  
FT DISULFID 394 521 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 432 448 BY SIMILARITY.  
FT DISULFID 440 510 BY SIMILARITY.  
FT DISULFID 535 604 BY SIMILARITY.  
FT DISULFID 567 583 BY SIMILARITY.  
FT DISULFID 594 622 BY SIMILARITY.  
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 492 492 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 546 546 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 644 644 R -> Q (IN REF. 2).  
SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B862ED7 CRC64;  
  
Query Match 32.0%; Score 735.5; DB 1; Length 655;  
Best Local Similarity 37.6%; Pred. No. 1e-50;  
Matches 160; Conservative 53; Mismatches 166; Indels 47; Gaps 9;  
  
QY 5 HQVPSNCDCLNGTGVSNKYFSNTHW-----CNCPKKFGQHCEDKSKTCYEGNGH 56  
Db 242 HTACLSPPCLMGHC-----HLIVATTTVCACPPGAGRLCNTEPDERCFLNGNT 292  
QY 57 FYRGKASTDTMGRECLPWN SATVLQOTYHAHRS DALQLGLGKHNVCNPNDRRRPWCYVQ 116  
Db 293 GYRGVASTSASGLSCLAWNSDLLYQELHVDVSGAALLGLGPHAYCRNPNDRPWCYV 352  
QY 117 VGLAPLVQECMVHDC-----ADGKKPSSPEELKFCQGX-----TLRPRFK 158  
Db 353 KDSALSWEYCRLEACESLTVQLSPDLLATLPEPASGRQ---ACGRHKRRTLRPR-- 407  
QY 159 IIGGEFTTIENQPFPAALYRRHSGSVYVCGSLISPCWVISATHCFIDYPKEDIVY 218  
Db 408 IIGSSSLPGSHPWLAII---IGDS---FCAGSLVHTCWVWSAAHCFSPSPDSVSV 461  
QY 219 IGRSLNSNTQGEKMFVEVNLILHKDYSADTLAHNDIALIKRSKRCQAQPSRTIOTI 278  
Db 462 LGQHFFNRTTDTVTQTGIEKVIPTLYSVFNPSDH-DLVLRLLKKGDRCATRSQFQPI 520  
QY 279 CLFSMYNDPQFGTSCETITGFGKENSTDYLPYQLKMTVVKLISHRECQPHYYGSEVTK 338

```
Db 521 CUPEPGSTPAGKQCLAGHLDENVGYSSLRALVPLVADHKCSSPEVYGADISPN 580
Qy 339 MLCAADPQWKTDSCQSDGSLVCSLQGRMTLGIIVSWGRGCALKDKPKGVYTRVSHFLPW 398
Db 581 MLCAGYFDCKSDACQSDGSLVCSLQGRMTLGIIVSWGRGCALKDKPKGVYTRVSHFLPW 640
Qy 399 IRSHTK 404
Db 641 INDIR 646

RESULT 17
HGFA_MOUSE
ID HGFA_MOUSE STANDARD; PRT; 653 AA.
AC Q9R098; Q9JKV4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
activator) (HGFA).
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Itoh H., Kataoka H., Koono H.;
RT "Mouse hepatocyte growth factor activator.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J.,
RT Yang J., Huan Y.;
RT "Activation of HGF by endogenous HGF activator is required for
metanephric kidney morphogenesis in vitro.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING
IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
DISULFIDE BOND (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY
SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement. (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
```

```
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SRR; 1.
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29 BY SIMILARITY.
FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
CHAIN.
FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
CHAIN.
FT DOMAIN 105 145 FIBRONECTIN TYPE-II.
FT DOMAIN 157 195 EGF-LIKE 1.
FT DOMAIN 197 237 FIBRONECTIN TYPE-I.
FT DOMAIN 238 276 EGF-LIKE 2.
FT DOMAIN 283 364 KRINGLE.
FT DOMAIN 406 653 SERINE PROTEASE.
FT ACT_SITE 445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 105 130 BY SIMILARITY.
FT DISULFID 119 145 BY SIMILARITY.
FT DISULFID 161 172 BY SIMILARITY.
FT DISULFID 166 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 199 227 BY SIMILARITY.
FT DISULFID 225 234 BY SIMILARITY.
FT DISULFID 242 253 BY SIMILARITY.
FT DISULFID 247 284 BY SIMILARITY.
FT DISULFID 286 275 BY SIMILARITY.
FT DISULFID 283 364 BY SIMILARITY.
FT DISULFID 304 346 BY SIMILARITY.
FT DISULFID 335 359 INTERCHAIN (BY SIMILARITY).
FT DISULFID 392 519 BY SIMILARITY.
FT DISULFID 430 446 BY SIMILARITY.
FT DISULFID 438 508 BY SIMILARITY.
FT DISULFID 533 602 BY SIMILARITY.
FT DISULFID 565 581 BY SIMILARITY.
FT DISULFID 592 620 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 544 544 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 164 164 G -> W (IN REF. 2).
SQ SEQUENCE 653 AA; 70567 MW; 88B4B20255DF7FDC CRC64;
```

Query Match 31.6%; Score 726; DB 1; Length 653;  
Best Local Similarity 37.3%; Pred. No. 5.6e-50;  
Matches 158; Conservative 57; Mismatches 167; Indels 42; Gaps 9;

Qy

5 HQVPSNCDCLNGTCVSNKYFSNIHW-----CNCPKFGGQHCEIDKSKTCYEGNGH 56



```
FT DISULFID 422 425 BY SIMILARITY.
FT DISULFID 488 557 BY SIMILARITY.
FT DISULFID 520 536 BY SIMILARITY.
FT DISULFID 547 578 BY SIMILARITY.
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66795 MW; 48DCB946FB9ED59 CRC64;

Query Match 31.2%; Score 718.5; DB 1; Length 603;
Best Local Similarity 36.8%; Pred. No. 2e-49;
Matches 161; Conservative 64; Mismatches 149; Indels 63; Gaps 13;

Qy 13 CLNGCTVSNKYFNSIHWCNCPKFGGCHCEIDSKSCYEGNGHYRGKASTDTMGRCPL 72
Db 182 CLNGRCLE---VEGHLLCDPCMGYTPGCDLDTASCYEGRGVSYRGMAITTVSGAKCQ 238

Qy 73 PWSATVLQOYTHAHRSD-ALQLGLGKHNYCRNPNRRRRPWCYVQVGLKPLVQSCWVHDC 131
Db 239 RWAS----EATYRNNTAEQALRRGLGHHTFCRNPNDTRPWCFFVMGNRLSWBYCDLAQC 294

Qy 132 ADGKKPSPPELKFQ-----CGOKTLRPR 156
Db 295 QYPPQPTATPHD-RFEHPKLPSSRLSILQPTQTOALANELPETSSLLCGQR-LRKR 352

Qy 157 F---KIIGBEFTTIENQFPAIYRRHRGVSIVYVCGSLISPCWVISATHCFIDYPKK 212
Db 353 LSSLRIYVGLVALPGAHPYIAALY----WGS--NFCGSLIAPCWVLTAAHCLQNRFP 406

Qy 213 EDYIVYGRSLNSNTQCEMKEFEVENLILHKDYADTLAHNDIALLKI-RSKEGRCAOP 271
Db 407 BELKVLQDQRHNSCEHCQTLAVHSYRLHBAFSPS--SYLNDLALLRLQKSDAGSCAOL 464

Qy 272 SRTIQTICLPSMYNDPQFG--TSCEITGFGKENSTDYLYPEQLKQVTVVKLISHRECOQPH 329
Db 465 SPYVTVCLPSGAPPSSETTCCEVAGCQHGFEGABEYSYFLQEAQVPLISSERCSSPE 524

Qy 330 YGSEVVTMLCAADPQWKTSCQDSGGPLVC---SLQGRMTLTGIVSWGRGCAALKDKP 386
Db 525 VHGDAFLGMLCAGLEGTTDACQDSGGPLVCDEAAEHLRLIRGVSMGSGGDRNKP 584

Qy 387 GYTVRVSHFLPWIRGHT 403
Db 585 GYITDVASILTVIQKHT 601

RESULT 19
ID FA12 HUMAN STANDARD; PRT; 615 AA.
AC P00748; P78339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor) (HAF).
OS F12.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88007593; PubMed=2888762;
RA Cool D.E., McGillivray R.T.A.;
RT "Characterization of the human blood coagulation factor XII gene. Intron/exon gene organization and analysis of the 5'-flanking region.";
RT J. Biol. Chem. 262:13662-13673 (1987).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
```

```
RN [3]
RP SEQUENCE OF 4-615 FROM N.A.
RX MEDLINE=86176794; PubMed=3754331;
RA Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA Cortese R.;
RT "cDNA sequence coding for human coagulation factor XII (Hageman).";
RN Nucleic Acids Res. 14:3146-3146(1986).
RN [4]
RP SEQUENCE OF 14-615 FROM N.A.
RX MEDLINE=86033930; PubMed=3877053;
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA McGallivray R.T.A.;
RT "Characterization of human blood coagulation factor XII cDNA. Prediction of the primary structure of factor XII and the tertiary structure of beta-factor Xlla.";
RN J. Biol. Chem. 260:13666-13676(1985).
RN [5]
RP SEQUENCE OF 146-615 FROM N.A.
RX MEDLINE=86216049; PubMed=3011063;
RA Que B.G., Davie E.W.;
RT "Characterization of a cDNA coding for human factor XII (Hageman factor).";
RN Biochemistry 25:1525-1528(1986).
RN [6]
RP SEQUENCE OF 20-379.
RX MEDLINE=85182674; PubMed=3886654;
RA McMullen B.A., Fujikawa K.;
RT "Amino acid sequence of the heavy chain of human alpha-factor Xlla (activated Hageman factor).";
RN J. Biol. Chem. 260:5328-5341(1985).
RN [7]
RP SEQUENCE OF 354-362 AND 373-615.
RX MEDLINE=83291041; PubMed=6604055;
RA Fujikawa K., McMullen B.A.;
RT "Amino acid sequence of human beta-factor Xlla.";
RN J. Biol. Chem. 258:10924-10933(1983).
RN [8]
RP SEQUENCE OF 561-615 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=96133302; PubMed=8528215;
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;
RT "The novel acceptor splice site mutation 11396(G->A) in the factor XII gene causes a truncated transcript in cross-reacting material negative patients.";
RN Hum. Mol. Genet. 4:1235-1237(1995).
RN [9]
RP CARBOHYDRATE-LINKAGE SITE THR-109.
RX MEDLINE=92184750; PubMed=1544894;
RA Harris R.J., Ling V.T., Spellman M.W.;
RT "O-linked fucose is present in the first epidermal growth factor domain of factor XII but not protein C.";
RN J. Biol. Chem. 267:5102-5107(1992).
RN [10]
RP VARIANT WASHINGTON D.C. SER-590.
RX MEDLINE=90046788; PubMed=2510163;
RA Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,
RA Saito H.;
RT "Coagulation factor XII (Hageman factor) Washington D.C.: inactive factor Xlla results from Cys-571-->Ser substitution.";
RN Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).
RN [11]
RP VARIANT LOCARNO PRO-372.
RX MEDLINE=94325559; PubMed=8049433;
RA Hovig J.K., Schaller J., Stricker H., Willemin W.A., Furlan M.,
RA Lammle B.;
RT "Coagulation factor XII Locarno: the functional defect is caused by the amino acid substitution Arg-353-->Pro leading to loss of a kallikrein cleavage site.";
RN Blood 84:1173-1181(1994).
RN [12]
RP VARIANT TENRI CYS-53.
RX MEDLINE=99290785; PubMed=10361128;
RA Kondo S., Tokunaga F., Kawano S., Oono Y., Kumagai S., Koide T.;
```



"Factor XII Tenri, a novel cross-reacting material negative factor XII deficiency, occurs through a proteasome-mediated degradation.";

-!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE GENERATION OF BRADYKININ AND ANGIOTENSIN.

-!- CATALYTIC ACTIVITY: Cleaves selectively Arg-Ile bonds in factor VII to form factor VIIa and factor XI to form factor XIa.

-!- PM: O- AND N-GLYCOSYLATED.

-!- DISEASE: DEFECTS IN F12 DO NOT CAUSE ANY CLINICAL SYMPTOMS. THE SOLE EFFECT IS THAT WHOLE-BLOOD CLOTTING TIME IS PROLONGED.

-!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY FACTOR XII TO FORM KALLIKEIN, WHICH THEN CLEAVES FACTOR XII FIRST TO ALPHA-FACTOR XIIa AND THEN TO BETA-FACTOR XIIa. ALPHA-FACTOR XIIa ACTIVATES FACTOR XI TO FACTOR XIa.

-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

-!- SIMILARITY: Contains 2 EGF-like domains.

-!- SIMILARITY: Contains 1 fibronectin type I domain.

-!- SIMILARITY: Contains 1 fibronectin type II domain.

-!- SIMILARITY: Contains 1 kringle domain.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; M31315; AAA70225.1; -.

EMBL; AF538691; AAM97932.1; -.

EMBL; M11723; AAA51986.1; -.

EMBL; M17466; AAB59490.1; -.

EMBL; M17464; AAB59490.1; JOINED.

EMBL; M17465; AAB59490.1; JOINED.

EMBL; M13147; AAA70224.1; -.

EMBL; U71274; AAB51203.1; -.

PIR; A29411; KFHU12.

HSSP; P00763; 1DPO.

MEROPS; S01.211; -.

Genew; HGNC:3530; F12.

MIN; 234000; -.

GO; GO:0003805; F: blood coagulation factor XI activity; TAS.

GO; GO:0003806; F: blood coagulation factor XII activity; TAS.

GO; GO:0007596; P: blood coagulation; TAS.

InterPro; IPR001314; Chymotrypsin.

InterPro; IPR008209; EGF-like.

InterPro; IPR000083; Fibrinctnl.

InterPro; IPR000562; FN Type\_II.

InterPro; IPR006210; IEGF.

InterPro; IPR000001; Kringle.

Pfam; PF00008; Ser\_Protease\_Try.

Pfam; PF00008; EGF\_2.

Pfam; PF00039; fn1; 1.

Pfam; PF00040; fn2; 1.

Pfam; PF00051; kringle; 1.

Pfam; PF00089; trypsin; 1.

PRINTS; PR00722; CHYMOTRYPSIN.

PRINTS; PR00013; FNTYPEII.

PRINTS; PR00018; KRINGLE.

ProDom; PD000995; FN Type\_II; 1.

ProDom; PD000395; Kringle; 1.

SMART; SM00181; EGF; 2.

SMART; SM00058; FN1; 1.

SMART; SM00059; FN2; 1.

SMART; SM00130; KR; 1.

SMART; SM00020; Tryp\_Spc; 1.

PROSITE; PS00022; EGF\_1; 2.

PROSITE; PS01186; EGF\_2; 1.

PROSITE; PS01253; FIBRONECTIN 1; 1.

PROSITE; PS00023; FIBRONECTIN 2; 1.

PROSITE; PS00021; KRINGLE\_1; 1.

DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;  
KW Hydrolase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;  
KW Polymorphism; Disease mutation.  
FT SIGNAL 1 19  
FT CHAIN 20 372 ALPHA-FACTOR XIIA HEAVY CHAIN.  
FT CHAIN 373 615 ALPHA-FACTOR XIIA LIGHT CHAIN.  
FT CHAIN 354 362 BETA-FACTOR XIIA PART 1.  
FT CHAIN 373 615 BETA-FACTOR XIIA PART 2.  
FT DOMAIN 47 88 FIBRONECTIN TYPE-II.  
FT DOMAIN 94 131 EGF-LIKE 1.  
FT DOMAIN 133 173 FIBRONECTIN TYPE-I.  
FT DOMAIN 174 210 EGF-LIKE 2.  
FT DOMAIN 217 295 KRINGLE.  
FT DOMAIN 296 349 PRO-RICH.  
FT DOMAIN 373 615 SERINE PROTEASE.  
FT CARBOHYD 109 109 O-LINKED (FUC).  
FT CARBOHYD 249 249 O-LINKED (GLCNAC...).  
FT CARBOHYD 299 299 O-LINKED (POTENTIAL).  
FT CARBOHYD 305 305 O-LINKED (POTENTIAL).  
FT CARBOHYD 308 308 O-LINKED (POTENTIAL).  
FT CARBOHYD 328 328 O-LINKED (POTENTIAL).  
FT CARBOHYD 329 329 O-LINKED (POTENTIAL).  
FT CARBOHYD 337 337 O-LINKED (POTENTIAL).  
FT ACT\_SITE 412 412 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 461 461 CHARGE RELAY SYSTEM (BY SIMILARITY).

Query Match 30.18; Score 692; DB 1; Length 615;

Best Local Similarity 34.3%; Pred. No. 2.6e-47;

Matches 153; Conservative 60; Mismatches 163; Indels 70; Gaps 9;

Qy 13 CLANGTCVSNKYFSNIHWCNCPKFGQGHCEIDSKTCYEGNGHFGKASTDMGRPCL 72

Db 183 CLHGRCLC---VEGRLHCHCPGVYGFDFDVTKASYDGRGLSYRLARTTUSGAPCQ 239

Qy 73 PWSATVLOQTY-HAHRSDALQLGLGHYCRNPNRRRPMWCYVQVGLKPLVQECVHDC 131

Db 240 PWAS-----EATYRNVTAEARNWGLGHAFCRNPDNDIRPWCFLNDRLSWEYCDLAQC 295

Qy 132 -----ADKKPSPEEL 144

Db 296 QTPTQAAPTFSRHLVPLMPAQPAPKPTTRTPPSQTPGALPAKEQPSLRTNG 355

Qy 145 KFCGQ---KTLRPRFKIIGGEFTTIENQFWFAIYRRHRGGSVTVYVCGSLISPCWVIS 201

Db 356 PLSCGQRLRKSLSMTFVVGVLALRGADHPYIAALYWGHS-----FCAGSLIAPCWILT 409

Qy 202 ATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKYASDTLAHNDIALKKI 261

Db 410 AAHCLQDRPAPEDLTVVVLGQERRNHSCEPCQTALAVRSYRLHEAFS--PVSQYHDLALLRL 467

Qy 262 R-SKEGRCAQPSRTIOTICLPSMYNDQFQTSCEITGFGKENSVDYLPQLKMTVVKLI 320

Db 468 QEDAGSCALLSPYQVCLPSGAARSETTLQVAGWGHQFEGAEYASFLQEAQPFPL 527

Qy 321 SHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQG---RMTLTGVSWG 377

Db 528 SLERCAPDVHGSSILPGMLCAGFLEGDTACQDGGPLVCEQDAERLTLQGIISWG 587

Qy 378 RGCALXDKPGVYTVRSVHFLPWIRSH 403

Db 588 SGCGRNKPVGVTDAVYLLAWIREHT 613

RESULT 20

FA12\_BOVIN

ID\_FAI2\_BOVIN STANDARD; PRT; 593 AA.

AC P98140;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)  
Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)  
(HAF) (Fragment).  
Bos taurus (Bovine).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
Bovidae; Bovinae; Bos.  
NCBI\_TaxID=9913;  
[1]  
SEQUENCE FROM N.A.  
TISSUE=Liver;  
MEDLINE=94242782; PubMed=8186251;  
RA Shibuya Y., Senba U., Okebe H., Kambara T., Yamamoto T.;  
RT "Primary structure of bovine Hageman factor (blood coagulation factor  
XII): comparison with human and guinea pig molecules.";  
RL Biochim. Biophys. Acta 1206:63-70(1994).  
[2]  
SEQUENCE OF 10-21; 350-364 AND 525-550.  
RX MEDLINE=77182112; PubMed=861210;  
RA Fujikawa K., Walsh A.K., Davie W.E.;  
RT "Isolation and characterization of bovine factor XII (Hageman  
factor).";  
RL Biochemistry 16:2270-2278(1977).  
CC -1- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN  
THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE  
GENERATION OF BRADYKININ AND ANGIOTENSIN.  
CC -1- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor  
VII to form factor VIIa and factor XI to form factor Xla.  
CC -1- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).  
CC -1- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HWM KININOGEN FORM A  
COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY  
FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST  
TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR  
XIIA ACTIVATES FACTOR XI TO FACTOR XIA. BOVINE FACTOR XII IS  
CLEAVED ONLY TO ALPHA-FACTOR XIIA AS IT LACKS THE TRYPSIN/  
KALLIKREIN CLEAVAGE SITE.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: Contains 2 EGF-like domains.  
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -1- SIMILARITY: Contains 1 fibronectin type II domain.  
CC -1- SIMILARITY: Contains 1 kringle domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL; S70164; AAB30804.2; -;  
PIR; S45281; S45281.  
DR HSSP; P00763; IDPO.  
DR MEROPS; S01.211; -;  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR000083; Fibnctnl.  
DR InterPro; IPR000562; FN\_Type\_II.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00008; EGF; 2.  
DR Pfam; PF00039; fn1; 1.  
DR Pfam; PF00040; fn2; 1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00013; FNTYPEII.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000995; FN\_Type\_II; 1.  
DR ProDom; PD000395; Kringle; 1.  
SMART; SM00181; EGF; 2.

DR SMART; SM00058; FN1; 1.  
DR SMART; SM00059; FN2; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 2.  
DR PROSITE; PS01185; EGF\_2; FALSE NEG.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00023; FIBRONECTIN\_2; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;  
Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 9 POTENTIAL.  
FT CHAIN 10 349 ALPHA-FACTOR XIIA HEAVY CHAIN.  
FT CHAIN 350 593 ALPHA-FACTOR XIIA LIGHT CHAIN.  
FT DOMAIN 37 78 FIBRONECTIN TYPE-II.  
FT DOMAIN 84 121 EGF-LIKE 1.  
FT DOMAIN 123 163 FIBRONECTIN TYPE-I.  
FT DOMAIN 164 200 EGF-LIKE 2.  
FT DOMAIN 207 287 KRINGLE.  
FT DOMAIN 297 333 PRO-RICH.  
FT DOMAIN 350 593 SERINE PROTEASE.  
FT ACT\_SITE 389 389 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 438 438 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 541 541 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 88 100 BY SIMILARITY.  
FT DISULFID 94 109 BY SIMILARITY.  
FT DISULFID 111 120 BY SIMILARITY.  
FT DISULFID 125 153 BY SIMILARITY.  
FT DISULFID 151 160 BY SIMILARITY.  
FT DISULFID 168 179 BY SIMILARITY.  
FT DISULFID 173 188 BY SIMILARITY.  
FT DISULFID 190 199 BY SIMILARITY.  
FT DISULFID 207 287 BY SIMILARITY.  
FT DISULFID 230 269 BY SIMILARITY.  
FT DISULFID 258 282 BY SIMILARITY.  
FT DISULFID 336 463 BY SIMILARITY.  
FT DISULFID 374 390 BY SIMILARITY.  
FT DISULFID 382 452 BY SIMILARITY.  
FT DISULFID 413 416 BY SIMILARITY.  
FT DISULFID 479 547 BY SIMILARITY.  
FT DISULFID 510 526 BY SIMILARITY.  
FT DISULFID 537 568 BY SIMILARITY.  
FT CARBOHYD 99 99 O-LINKED (FUC) (BY SIMILARITY).  
FT CARBOHYD 241 241 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 263 263 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 593 AA; 65148 MW; 721592BA792BD61F CRC64;  
Query Match 27.9%; Score 642; DB 1; Length 593;  
Best Local Similarity 35.1%; Pred.No.2.2e-43;  
Matches 155; Conservative 60; Mismatches 167; Indels 60; Gaps 15;  
QY 6 QVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYE--GNGHFYRGKAS 63  
Db 166 QVCRTNPCLNGDSCLQAE--GHRLCRCAPSPAGRLCDVDLKASCYDDRRDGLSYRMAG 222  
QY 64 TDTMGPPCLPWN SATVLOQTY-HAHRSDALQLGLGHKNYCRPNDRRRRWCVYVQGLKPL 122  
Db 223 TTLSGAPCQSWAS----EATYWNVTAEQVLNGLGDLGHAFCRPNPDNTRPWCFTWKGDRLS 278  
QY 123 VQECMVHDC-ADG-----KKPSS-----PPELKF-----CCGQ 150  
Db 279 WNYCLAPCAAAGHEHFFLPSPALQKDPSTTQTLPSLTSGWCSPTPLASGGPGCGQ 338  
QY 151 ----KTLRPFKLIIGFEFTTIEQNQPFALYRRHGGSVTVYCGGLSPICWVISATHCFI 207  
Db 339 RLRLKWLSSILNRVVGVLVALPGAHPYIALYWDQ-----HFCAGSLIAPCWLTAAHCLQ 392

Qy 208 DYPKEDYVYVGRSLNNTQGMKFEVENILHKOYSAOTLAHNDIALKIR-SKEG 266  
Db 393 NRPAPKELTVVLGQDRHNSCEQCOTLAIRDYRLHEAFPTYQH--DLALVRLQESADG 450  
Qy 267 RCAOPSTIOTICLPMSYNDPQFGS--CELTGKKNSTDYLPQELKMTVVKLISHRE 324  
Db 451 CCAHPSFPVQVCLPSTAAAESEAAVCEVAGHGHQFEGGE--YSSFLQEAQVPLIDPQR 509  
Qy 325 CQPHYGYGVETTKMLCAADPWKTDSCQDSSGGPLVC---SLQRMFLTGIVSWGRGCA 381  
Db 510 CSAFDVHGAAFTQGMCLGAFLEGGTDACQDSSGGPLVCDETPERQLILRGIVSWSGCG 569  
Qy 382 LKDKPGVYTRVSHFLPWRSHST 403  
Db 570 NRLKPGVYTDVANYLAMIREHT 591

RESULT 21  
PLMN\_PIG STANDARD; PRT; 790 AA.  
AC P06867;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Plasminogen (EC 3.4.21.7).  
CN PLG.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE OF 1-560.  
RA Schaller J., Marti T., Roessellet S.J., Kaempfer U., Rickli E.B.;  
RT "Amino acid sequence of the heavy chain of porcine plasmin. Comparison  
of the carbohydrate attachment sites with the human and bovine  
species.";  
RL Fibrinolysis 1:91-102(1987).  
RP [2]  
RP SEQUENCE OF 450-790.  
RA MEDLINE=85203907; PubMed=3846533;  
RM Marti T., Schaller J., Rickli E.B.;  
RT "Determination of the complete amino-acid sequence of porcine  
miniplasminogen.";  
RN Eur. J. Biochem. 149:279-285 (1985).  
RP [3]  
RP CARBOHYDRATE-LINKAGE SITES.  
RA MEDLINE=88185329; PubMed=3356193;  
RM Marti T., Schaller J., Rickli E.B., Schmid K., Kamerling J.P.,  
RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;  
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine  
plasminogen. Species specificity in relation to sialylation and  
fucosylation patterns.";  
RL Eur. J. Biochem. 173:57-63(1988).  
CC -|- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS  
A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING  
EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,  
AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE  
GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN  
ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH  
AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,  
LAMININ AND VON WILLEBRAND FACTOR.  
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;  
higher selectivity than trypsin. Converts fibrin into soluble  
products.  
CC -|- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN  
ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO  
FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.  
CC -|- PTM: N-LINKED GLYCAN CONTAINS N-ACETYLACTOSAMINE, SIALIC ACID AND  
IS CORE FUCOSYLATED. O-LINKED GLYCANS CONSIST OF GAL-GALNAc  
DISACCHARIDE WITH IS MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES  
(MICROHETEROGENEITY).  
CC -|- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN  
IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.

CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.  
CC -|- SIMILARITY: Contains 5 kringle domains.  
DR PIR: S03733; PLRG.  
DR HSSP: P00747; SHPG.  
DR MEROPS: S01.233; --  
DR GlycoSuiteDB: P06867; --  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR000001; Kringle.  
DR InterPro: IPR003014; PAN.  
DR InterPro: IPR003609; Pan\_app.  
DR InterPro: IPR003666; Prothrombin.  
DR InterPro: IPR001254; Ser\_protease\_Try.  
DR Pfam: PF00051; kringle; 5.  
DR Pfam: PF00024; PAN; 1.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR PRINTS: PR00018; KRINGLE.  
DR PRINTS: PR01505; PROTHROMBIN.  
DR PRODOM: PD000395; Kringle; 5.  
DR SMART: SM00130; KR; 5.  
DR SMART: SM00473; PAN AP; 1.  
DR SMART: SM00020; Tryp\_SPC; 1.  
DR PROSITE: PS00021; KRINGLE\_1; 5.  
DR PROSITE: PS00070; KRINGLE\_2; 5.  
DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.  
FT CHAIN 1 560  
FT CHAIN PLASMIN HEAVY CHAIN A.  
FT DOMAIN 561 790  
FT DOMAIN SERINE PROTEASE.  
FT DOMAIN 84 162  
FT DOMAIN KRINGLE 1.  
FT DOMAIN 166 243  
FT DOMAIN KRINGLE 2.  
FT DOMAIN 256 333  
FT DOMAIN KRINGLE 3.  
FT DOMAIN 358 435  
FT DOMAIN KRINGLE 4.  
FT DOMAIN 461 540  
FT DOMAIN KRINGLE 5.  
FT ACT SITE 602 602  
FT ACT SITE CHARGE RELAY SYSTEM.  
FT ACT SITE 645 645  
FT ACT SITE CHARGE RELAY SYSTEM.  
FT ACT SITE 740 740  
FT CARBOHYD N-LINKED (GLCNAC...).  
FT CARBOHYD /FTID=CAR\_000019.  
FT CARBOHYD O-LINKED (GALNAc...).  
FT /FTID=CAR\_000020.  
SQ SEQUENCE 790 AA; 88592 MW; F04EA06E74BCD58E CRC64;  
Query Match 21.98; Score 504.5; DB 1; Length 790;  
Best Local Similarity 35.48; Pred. No. 2.3e-32;  
Matches 130; Conservative 45; Mismatches 141; Indels 51; Gaps 14;  
Qy 45 DSKTCYGVGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHR---SDALQLGLGKH 100  
Db 456 DLSEDCMFGNGKRYGKRAITVAGVPCQEWAA-----QEPHRSITFTETNPAGLEK-N 509  
Qy 101 YCRNPD-NRRRPNVCYQVQGLKPLVQECMVHDCADGKPPSPBELKPCQCKTLRPR--- 156  
Db 510 YCRNPDGDDNGPWCYT-TNPQKLFDCYDVPQCVTS-----SFDCKGKVPKPKCP 558  
Qy 157 FKIGGEFTTIENQPFPAALYRHRGSSVYVCGSGLISPCWVISATHCFIDYPKKEDIY 216  
Db 559 ARVVGCVSIPHPWQISLURYRG-----HFCGGLISPEWVLTAKHLEKSSSSPSYK 614  
Qy 217 VYLGSRSLNSNTQGMKFEVENILHKOYSAOTLAHNDIALKIRSKGRCQAPSRITQ 276  
Db 615 VILGAHEEYHLGSGVQEIDVSKLF--KEPS-----EADIALKLSSP-----AVTDKVI 662  
Qy 277 TICLPWYNDPQF-----GTSCEITGKGNSTDYLPQELKMTVVKLISHRECOQPHYVG 332  
Db 663 PACLPF-----PNYVADRTACTYTGWGETKGT--YGAGLLKEARLPVIEKNVCNRYEYL 716  
Qy 333 SEVYTRMLCAADPQWKTDCQDSSGGPLVCSLQGRMTLTGIVSWGRGCAIKQKPGYVTRV 392  
Db 717 GKVSFNLGAGHLAGGIDSCQDSSGGPLVCFERDKYILQGVTSWGLGCAIPKPKGVYVRV 776

QY 393 SHELPMI 399  
 Db 777 SRFVTWI 783  
 RESULT 22  
 ID PLMN\_MOUSE STANDARD; PRT; 812 AA.  
 AC F20918; Q8CIS2; Q91WJ5;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].  
 GN PLG.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91184812; PubMed=2081600;  
 RA Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;  
 RT "Characterization of the cDNA coding for mouse plasminogen and  
 RT localization of the gene to mouse chromosome 17";  
 RL Genomics 8:49-61(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/Sv;  
 RA Brathwaite M., Waeltz P., Qian Y., Dudekula D., Schlessinger D.,  
 RA Nagaraja R.;  
 RT "Genomic sequence analysis in the mouse t-complex region";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustun T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 1-16 FROM N.A.  
 RC STRAIN=129/SvJ; TISSUE=Liver;  
 RX PubMed=12149246;  
 RA Bannach F.G., Gutierrez A., Fowler B.J., Bugge T.H., Degen J.L.,  
 RA Farmer R.J., Miles L.A.;  
 RT "Localization of regulatory elements mediating constitutive and  
 RT cytokine-stimulated plasminogen gene expression.";  
 RL J. Biol. Chem. 277:38579-38588(2002).  
 RN [5]  
 RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.  
 RX MEDLINE=95042728; PubMed=7525077;  
 RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,  
 RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;  
 RT "Angiostatin: a novel angiogenesis inhibitor that mediates the  
 RT suppression of metastases by a Lewis lung carcinoma.";

RL Cell 79:315-328(1994).  
 CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS  
 CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING  
 CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,  
 CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE  
 CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UKONINASE-TYPE PLASMINOGEN  
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH  
 CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,  
 CC LAMININ AND VON WILLEBRAND FACTOR.  
 CC -!- FUNCTION: ANGIOSTATIN IS AN ANGIOGENESIS INHIBITOR THAT BLOCKS  
 CC NEOVASCULARIZATION AND GROWTH OF EXPERIMENTAL PRIMARY AND  
 CC METASTATIC TUMORS. IN VIVO.  
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;  
 CC higher selectivity than trypsin. Converts fibrin into soluble  
 CC products.  
 CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN  
 CC ACTIVATORS. BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO  
 CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.  
 CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN  
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.  
 CC -!- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION  
 CC INVOLVES ONLY CLEAVAGE AFTER ARG-581, RESULTING IN 2 CHAINS HELD  
 CC TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE  
 CC ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.  
 CC -!- SIMILARITY: Contains 5 kringle domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J04766; AAA50168.1; -;  
 CC EMBL; AF481053; AAM22156.1; -;  
 CC EMBL; BC014773; AAH14773.1; -;  
 CC EMBL; AY134430; AAN15805.1; -;  
 CC FIR; A38514; PLMS.  
 CC HSSP; P00747; 1PMK.  
 CC MEROPS; S01.233; -;  
 CC MGD; MGI:97620; Pig.  
 CC GO; GO:0016506; P:apoptosis activator activity; IDA.  
 CC GO; GO:0006915; P:apoptosis; IDA.  
 CC InterPro; IPR001314; Chymotrypsin.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR003014; PAN.  
 CC InterPro; IPR003609; Pan\_app.  
 CC InterPro; IPR003966; Prothrombin.  
 CC InterPro; IPR001254; Ser\_protease\_Try.  
 CC Pfam; PF00051; kringle; 5.  
 CC Pfam; PF00024; PAN; 1.  
 CC Pfam; PF00089; trypsin; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC PRINTS; PR00018; KRINGLE.  
 CC PRINTS; PR01505; PROTHROMBIN.  
 CC ProDom; PD000395; Kringle; 5.  
 CC SMART; SM00130; KR; 5.  
 CC SMART; SM00473; PAN\_AP; 1.  
 CC SMART; SM00020; Tryp\_Spc; 1.  
 CC PROSITE; PS00021; KRINGLE\_1; 5.  
 CC PROSITE; PS00070; KRINGLE\_2; 5.  
 CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
 CC Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;  
 CC Signal.  
 CC SIGNAL. 1 19  
 CC CHAIN 20 812 PLASMINOGEN.  
 CC CHAIN 20 581 PLASMIN HEAVY CHAIN A.  
 CC PEPTIDE 20 97 ACTIVATION PEPTIDE.

```
FT CHAIN 98 581 PLASMIN SHORT FORM OF CHAIN A.
FT CHAIN 98 7436 ANGIOTATIN.
FT CHAIN 582 812 PLASMIN LIGHT CHAIN B.
FT DOMAIN 103 181 KRINGLE 1.
FT DOMAIN 184 262 KRINGLE 2.
FT DOMAIN 275 352 KRINGLE 3.
FT DOMAIN 377 454 KRINGLE 4.
FT DOMAIN 481 560 KRINGLE 5.
FT DOMAIN 582 812 SERINE PROTEASE.
FT ACT_SITE 667 667 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 762 762 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 49 73 BY SIMILARITY.
FT DISULFID 53 61 BY SIMILARITY.
FT DISULFID 103 181 BY SIMILARITY.
FT DISULFID 124 164 BY SIMILARITY.
FT DISULFID 152 176 BY SIMILARITY.
FT DISULFID 185 262 BY SIMILARITY.
FT DISULFID 188 316 BY SIMILARITY.
FT DISULFID 206 245 BY SIMILARITY.
FT DISULFID 234 257 BY SIMILARITY.
FT DISULFID 275 352 BY SIMILARITY.
FT DISULFID 324 347 BY SIMILARITY.
FT DISULFID 377 454 BY SIMILARITY.
FT DISULFID 426 449 BY SIMILARITY.
FT DISULFID 481 560 BY SIMILARITY.
FT DISULFID 502 543 BY SIMILARITY.
FT DISULFID 531 555 BY SIMILARITY.
FT DISULFID 568 687 INTERCHAIN (BY SIMILARITY).
FT DISULFID 578 586 INTERCHAIN (BY SIMILARITY).
FT DISULFID 609 625 BY SIMILARITY.
FT DISULFID 701 768 BY SIMILARITY.
FT DISULFID 731 747 BY SIMILARITY.
FT DISULFID 758 786 BY SIMILARITY.
FT CONFLICT 235 235 R -> H (IN REF. 1).
FT CONFLICT 525 525 G -> D (IN REF. 1).
FT CONFLICT 649 649 S -> L (IN REF. 1).
FT CONFLICT 649 649 S -> L (IN REF. 1).
SQ SEQUENCE 812 AA; 90781 MW; 2417326056A2FFD2 CRC64;

Query Match
Best Local Similarity 21.8%; Score 501; DB 1; Length 812;
Matches 135; Conservative 48; Mismatches 170; Indels 62; Gaps 15;

QY 10 NCDCLNGTGVSNKYFNHWCNPK--KFGQHCIEI-----DKSKTCYENGH 56
Dy 428 NPDGDKGPWCYTTDPVSRWEYCNLRCSETGSGVVELFTVSGEPSPDSSETDCMYGNK 487
QY 57 FYRKASTDTRGRCLPWNSATVLQQTVAHRSDALQ----LGLGKKNYCNRPD-NRRRP 111
Dy 488 DYRGKTAVTAAATGTCQGWAA-----QPHRHSIFPTQTNFAGLEK-NYCNRPDGVN 541
QY 112 WCYVQVGLKPLVQCNVHDCADGKKSPPELKFQCGQKTLRPR---FKIIGGEFTTIE 168
Dy 542 WCYT-TNPRKLYDCDIPLCASAS-----SFEGCKPQVEPKPCGRVVGCVANPH 591
QY 169 NQPFALYRHRGSGVYVCGGSLSPCWVISATHCFIDYPKEDIYVILGRSLNSNT 228
Dy 592 SWPQISLIRFTGT---QHFCGGTLIAPEWVLTAAHCLKSSRPFYKVLGAHEEYIRG 648
QY 229 QGEMKFEVENLHKDYASDTLAHNDIALKIRSKRCRCAQPSRTQICLPSMYNDPQ 288
Dy 649 SDVQEISVAKLLE-----PNNRDIALKL-----SRPATITDKVIPACLPS----PN 692
QY 289 F-----GTSCEITGKGNSTLYLPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAAD 344
Dy 693 YMWADRTICVTNGETQGT---FGAGRLKEAQLPVIENKVCNRVEYLNLRVKSTELCAGQ 750
QY 345 PQWKTDCQDGGGLVCSLQGRMTLTGIVSGWGGCALKOKPGYVTVSHPLPW 399
Dy 751 LAGVDSQDGGGLVCFEAKDKVILQGVTSWGLGRCARPNKPGYVTVSRFVDWI 805
```

```
RESULT 23
PLMN HUMAN
ID PLMN HUMAN STANDARD; PRT; 810 AA.
AC P00747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318848;
RA Petersen T.E., Maritzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
RL in the fibrinolytic system.";
RN J. Biol. Chem. 265:6104-6111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
RL for human plasminogen.";
RN FEBS Lett. 213:254-260(1987).
RN [3]
RP SEQUENCE OF 20-810.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the FIR data bank.
RN [4]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RL human and bovine plasminogen.";
RN Biochemistry 23:4243-4250(1984).
RN [5]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
RL of human plasminogen and their interaction with the NH2-terminal
RT activation peptide as studied by affinity chromatography.";
RN Eur. J. Biochem. 50:489-494(1975).
RN [6]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
RA Sottrup-Jensen L., Claess H., Zajdel M., Petersen T.E., Magnusson S.;
RL (In) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
RT Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
RL Raven Press, New York (1978).
RN [7]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide fragment from human
RL plasminogen that forms the linkage between the plasmin chains.";
RN Eur. J. Biochem. 58:539-547(1975).
RN [8]
RP SEQUENCE OF 581-810.
RX MEDLINE=7725245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RN Eur. J. Biochem. 76:129-137(1977).
RN [9]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
RL of human plasmin: light (B) chain active center histidine sequence.";
RN J. Biol. Chem. 248:1631-1633(1973).
```

RT RX MEDLINE=69234739; PubMed=4240117;  
RA Grobkopf W.R., Summaria L., Robbins K.C.;  
RT "Studies on the active center of human plasmin. Partial amino acid  
RT sequence of a peptide containing the active center serine residue.";  
RL J. Biol. Chem. 244:3590-3597(1969).  
[11]  
RN OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.  
RX MEDLINE=82213905; PubMed=6919539;  
RA Trexler M., Vali Z., Patthy L.;  
RT "Structure of the omega-aminocarboxylic acid-binding sites of human  
RT plasminogen. Arginine 70 and aspartic acid 56 are essential for  
RT binding of ligand by kringle 4.";  
RL J. Biol. Chem. 257:7401-7406(1982).  
[12]  
RN FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.  
RX MEDLINE=85054794; PubMed=6094526;  
RA Vali Z., Patthy L.;  
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34  
RT are essential for fibrin affinity of the kringle 1 domain.";  
RL J. Biol. Chem. 259:13690-13694(1984).  
[13]  
RN PHOSPHORYLATION SITE SER-597.  
RX MEDLINE=97345939; PubMed=9201958;  
RA Wang H., Prorok M., Bretthauer R.K., Castellino F.J.;  
RT "Serine-578 is a major phosphorylation locus in human plasma  
RT plasminogen.";  
RL Biochemistry 36:8100-8106(1997).  
[14]  
RN CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=88185329; PubMed=3356193;  
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,  
RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;  
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine  
RT plasminogen. Species specificity in relation to sialylation and  
RT fucosylation patterns.";  
RL Eur. J. Biochem. 173:57-63(1988).  
[15]  
RN "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of  
RT human plasminogen 2.";  
RL J. Biol. Chem. 272:7408-7411(1997).  
[16]  
RN CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.  
RX MEDLINE=95042728; PubMed=7525077;  
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,  
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;  
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the  
RT suppression of metastases by a Lewis lung carcinoma.";  
RL Cell 79:1315-1328(1994).  
[17]  
RN CHARACTERIZATION OF ANGIOSTATIN.  
RX MEDLINE=97238710; PubMed=9102221;  
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,  
RA Lapevich R., Nacy C.A.;  
RT "A recombinant human angiostatin protein inhibits experimental primary  
RT and metastatic cancer.";  
RL Cancer Res. 57:1329-1334(1997).  
[18]  
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.  
RX MEDLINE=92031502; PubMed=1657148;  
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;  
RT "Crystal and molecular structure of human plasminogen kringle 4  
RT refined at 1.9-A resolution.";  
RL Biochemistry 30:10576-10588(1991).  
[19]  
RN X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.  
RX MEDLINE=92031503; PubMed=1657149;  
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;

RT "The refined structure of the epsilon-aminocaproic acid complex of  
RT human plasminogen kringle 4.";  
RL Biochemistry 30:10589-10594(1991).  
[20]  
RN X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.  
RA Stec B., Yamano A., Whitlow M., Teeter M.M.;  
RT "Structure of human plasminogen kringle 4 at 1.68 Angstrom and 277 K.  
RT A possible structural role of disordered residues.";  
RL Acta Crystallogr. D 53:169-178(1997).  
[21]  
RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.  
RX MEDLINE=96180681; PubMed=8611560;  
RA Mathews I.I., Vanderhoff-Hanover P., Castellino F.J., Tulinsky A.;  
RT "Crystal structures of the recombinant kringle 1 domain of human  
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid  
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";  
RL Biochemistry 35:2567-2576(1996).  
[22]  
RN X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.  
RX MEDLINE=98198034; PubMed=9521645;  
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,  
RA Castellino F.J.;  
RT "Structure and ligand binding determinants of the recombinant kringle  
RT 5 domain of human plasminogen.";  
RL Biochemistry 37:3258-3271(1998).  
[23]  
RN STRUCTURE BY NMR OF 96-184.  
RX MEDLINE=94237157; PubMed=8181475;  
RA Rejante M.R., Llinas M.;  
RT "1H-NMR assignments and secondary structure of human plasminogen  
RT kringle 1.";  
RL Eur. J. Biochem. 221:927-937(1994).  
[24]  
RN STRUCTURE BY NMR OF 96-184.  
RX MEDLINE=94237158; PubMed=8181476;  
RA Rejante M.R., Llinas M.;  
RT "Solution structure of the epsilon-aminohexanoic acid complex of  
RT human plasminogen kringle 1.";  
RL Eur. J. Biochem. 221:939-949(1994).  
[25]  
RN STRUCTURE BY NMR OF 183-354.  
RX MEDLINE=96194156; PubMed=8652577;  
RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,  
RA Rickli E.E.;  
RT "Recombinant gene expression and 1H NMR characteristics of the  
RT kringle (2 + 3) supermodule: spectroscopic/functional individuality  
RT of plasminogen kringle domains.";  
RL Biochemistry 35:2357-2364(1996).  
[26]  
RN STRUCTURE BY NMR OF 374-461.  
RX MEDLINE=90219023; PubMed=2157850;  
RA Atkinson R.A., Williams R.J.P.;  
RT "Solution structure of the kringle 4 domain from human plasminogen by  
RT 1H nuclear magnetic resonance spectroscopy and distance geometry.";  
RL J. Mol. Biol. 212:541-552(1990).  
[27]  
RN VARIANTS PHE-374 AND THR-620.  
RT Query Match 21.7%; Score 499; DB 1; Length 810;  
RT Best Local Similarity 33.5%; Pred. No. 6.4e-32;  
RT Matches 139; Conservative 44; Mismatches 168; Indels 64; Gaps 16;

QY 10 NCDLNGGTCVSNKYFSNIHWCNCPKFGGQ-----HCEIDKSKTCYEGNGH 56  
DB 428 NPDADKGPWCFTTDPVSWEYCNLKKCSGTEASVAPPVLLPDVETSEEDCFNGCK 487  
QY 57 FYRGKASTDTMGRPCLPWNSATVLOQTYAHR-----SDALQLGLGHNYCRNPD-NRRRP 111  
DB 488 GYRGKRAITVTGTGTCQDWA-----QEPHRHSIFTFETNPRAGLEK-NYCRNPDGVDGP 541  
QY 112 WCYVQVGLKPLVQECWVHDCADKKPSSPELKFQCGQKTLRPR---FKITGGFTTIE 168  
DB 542 WCYT-TNPRKLYDYCDVQPCA---APS-----FDCGKPQVEPKCGRVVVGCVVAHPH 590



Query Match 21.6%; Score 497; DB 1; Length 810;  
Best Local Similarity 33.7%; Pred. No. 9.2e-32;  
Matches 140; Conservative 46; Mismatches 165; Indels 64; Gaps 16;

QY 10 NCCLNGGTCVSNKYFSNIHWCNCPKFKGQ-----HCEIDSKSKTCYEGNGH 56  
Db 428 NPADKGPWCFTDPDSVRWEYCNLKCSTEGSVAAPPVVAQLPDAETSEEDCMFGNGK 487

QY 57 FYRGKASTDTMGRPCLPWNSATVLOQTYHAHR-----SDALQLGLGKHNYCRNP 111  
Db 488 GYRGKATTVTGPCCQEWAA-----QEPHSHRIFTPTNPAGLEK-NYCRNPDGVDGVP 541

QY 112 WCYVQVGLKPLVQECMVHDCADGKSPPEELKFCGQKTLPR---FKIIGETTIE 168  
Db 542 WCYT-TNPKRLFDYCDVPOCAAS-----SFDGKRPQVPEKPCGRVVGCVAYPH 590

QY 169 NOPWFAIYRRHRGGSVTVCVCGSLSPCWVISATHCFIDYPKKDYIVVLGRSLNSNT 228  
Db 591 SWPW--QISLRTELG--MHFCGGTLISPEWLTAAHCLSKSRPSFYKVLGNHR-----641

QY 229 QGEMKEVENLILHKDYSADTLAHNDIALKIRSKGRCAPSRTIQTICLPSMYNDPQ 288  
Db 642 EVHLEPHVQEIYSKMFSEPARA---DIALKLSSP---AIIIDKVIAPCLPS-----DN 690

QY 289 F-----GTSCEITGFGKENSTDYLPQLKMTVVKLISHRECOPHYVGSVVTMCLCAAD 344  
Db 691 YVVADETECFITGWGTQGT--YGAGLLKEARLPVLENKVCNRYEFLNGVTVKITELCAGH 748

QY 345 POWKTDSCQDGGPLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPMI 399  
Db 749 LAGGTDSQDGGPLVCFEKDKYILQGVTSWGLGCARPKNKPGVYVRSRFTWI 803

RESULT 25  
PLMN\_CANFA STANDARD; PRT; 333 AA.

AC P80009;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Plasminogen (BC 3.4.21.7) (Fragment).  
GN PLG.

OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE.

RC TISSUE=Plasma;  
RX MEDLINE=90175323; PubMed=2626424;  
RA Schaller J., Straub C., Kaempfer U., Rickli E.E.;  
RT "Complete amino acid sequence of canine miniplasminogen.";  
RL Protein Seq. Data Anal. 2:445-450(1989).  
CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS  
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING  
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,  
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE  
CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN  
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH  
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,  
CC LAMININ AND VON WILLEBRAND FACTOR.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;  
CC higher selectivity than trypsin. Converts fibrin into soluble  
CC products.  
CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN  
CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO  
CC FIBRIN. ACTIVATED WITH UROKINASE AND HIGH CONCENTRATIONS OF  
CC STREPTOKINASE.  
CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN  
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.  
CC -!- SIMILARITY: Contains at least 1 kringle domain.

DR HSP; P00747; SHPG.  
DR MEROPS; S01.233; -.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003966; Prothrombin.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00089; trypsin; 1.  
DR Pfam; PF00051; kringle; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR0018; KRINGLE.  
DR PRINTS; PR01505; PROTHROMBIN.  
DR ProDom; PD000355; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp SPC; 1.  
DR PROSITE; PS00021; KRINGLE 1; 1.  
DR PROSITE; PS00070; KRINGLE 2; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen.  
FT NON\_TER 1 1  
FT CHAIN <1 103 PLASMIN HEAVY CHAIN A.  
FT CHAIN 104 333 PLASMIN LIGHT CHAIN B.  
FT DOMAIN 4 333 KRINGLE 5.  
FT DOMAIN 104 333 SERINE PROTEASE.  
FT DISULFID 4 83 BY SIMILARITY.  
FT DISULFID 25 66 BY SIMILARITY.  
FT DISULFID 54 78 BY SIMILARITY.  
FT DISULFID 90 208 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 100 108 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 130 146 BY SIMILARITY.  
FT DISULFID 222 289 BY SIMILARITY.  
FT DISULFID 252 268 BY SIMILARITY.  
FT DISULFID 279 307 BY SIMILARITY.  
FT ACT\_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 188 188 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 283 283 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT SITE 152 152 STREPTOKINASE-BINDING SITE (PROBABLE).  
FT SITE 186 186 STREPTOKINASE-BINDING SITE (PROBABLE).  
FT SITE 264 264 STREPTOKINASE-BINDING SITE (PROBABLE).  
FT SITE 277 277 SITE OF SUBSTRATE SPECIFICITY (BY SIMILARITY).  
SQ SEQUENCE 333 AA; 36678 MW; C8C0271B6C6AC8D4 CRC64;

Query Match 21.5%; Score 494.5; DB 1; Length 333;  
Best Local Similarity 34.8%; Pred. No. 5.3e-32;  
Matches 126; Conservative 49; Mismatches 136; Indels 51; Gaps 14;

QY 50 CYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHR-----SDALQLGLGKHNYCRNP 105  
Db 4 CMFGNGKGYRGKATTVMGIPQEWAA-----QEPHSHRIFTPTNPQAGLEK-NYCRNP 57

QY 106 D-NRRRPPWCYCVQVGLKPLVQECMVHDCADGKSPPEELKFCGQKTLPR---FKIIG 161  
Db 58 DGDVNGPWCYT-MNQRKLFDYCDVQCV-----STSDCKRPQVPEKPCGRVVG 106

QY 162 GEFTTIENQPFPAALYRRHRGGSVTVCVCGSLSPCWVISATHCFIDYPKKEDYIVLGR 221  
Db 107 GCVANPHSPWQISLRTRYG---KHFCGGTLISPEWLTAAHCLERSRSPASYKVLG- 161

QY 222 SELNSNTGEMKFVENLILHKDYSADTLAHNDIALKIRSKGRCAPSRTIQTICLP 281  
Db 162 ----AHKEVNLESVDQIEVYKLFLEPTRA---DIALKLSSP-----AVITSKVIPACL 210

QY 282 SWYNDPQ-----GTSCEITGFGKENSTDYLPQLKMTVVKLISHRECQPHYVSEVTT 337  
Db 211 ----PPNVVADRTLICYITGNETQGT--YGAGLLKEARLPVLENKVCNRYEFLNGRVKS 264

QY 338 KMLCAADPQWKTSCQDGGPLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLP 397  
Db 265 TELCAGNLAGTDSQDGGPLVCFEKDKYILQGVTSWGLGCARPKNKPGVYVRSRFT 324



```

QY 398 WI 399
  ||
Db 325 WI 326

RESULT 26
ID PLMN SHEEP STANDARD; PRT; 343 AA.
AC P81286;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragment).
GN PLG.
OS Ovis aries (sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP MEDLINE=93149995; PubMed=1492092;
RX Schaller J., Straub C., Kampfer U., Rickli E.E.;
RT "Complete amino acid sequence of ovine miniplasminogen.";
RL Protein Seq. Data Anal. 5:21-25(1992).
CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS. BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: Contains at least 2 kringle domains.
DR PIR; B61545; B61545.
DR HSP; P00747; SHPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.
FT NON_TER 1 1
FT DOMAIN <1 140 HEAVY CHAIN A.
FT DOMAIN 141 >343 LIGHT CHAIN A.
FT DOMAIN <1 17 KRINGLE 4.
FT DOMAIN 41 120 KRINGLE 5.
FT DOMAIN 114 341 SERINE PROTEASE.
FT ACT_SITE 181 181 CHARGE RELAY SYSTEM.
FT ACT_SITE 224 224 CHARGE RELAY SYSTEM.

FT ACT_SITE 319 319 CHARGE RELAY SYSTEM.
FT NON_TER 343 343
SQ SEQUENCE 343 AA; 37662 MW; 8DF6EBA92D596E0 CRC64;

Query Match 21.4%; Score 493; DB 1; Length 343;
Best local similarity 35.1%; Pred. No. 7.2e-32;
Matches 127; Conservative 41; Mismatches 142; Indels 52; Gaps 13;

QY 50 CYEGNGHYFGKASTDTMGPCPLPWNASATVLQTYHAH---RSDALQLGLGKHCNRP 105
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15 CMIGIGKGYGKATTVAGVPCQEWAA-----QEPHGHGIFTETNFRAGLEK-NYCRNP 68
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 106 D-NRRPWCYVQVLKPLVQECMVHDCADGKSPPEELKFCQGQKTLRPR---FKIIG 161
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69 DGDVNGWCYT-TNPRKLFDCIPQC-----ESSFDCGKPKVKKPCARVVG 116
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 162 GERTTIENQPFPAAIRHRHGGSVTVVCGSLSPCWVISATHCFDIYPKKEDIVYVLR 221
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 117 GCVATPHSWPQVSLRRSR-----EHFCGGTLISPEWVLTAAHCLDSILGSPFYVILGA 172
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 222 SRLNSNTQGMKFEVENLIILHKDYSADTLAHHNDIALLKIRSKRGCAQPSRTIQTICLP 281
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 HYEMAREASQVEIPVSEFLFLEPSRA-----DIALKLSSP-----AVITDEVIPACLP 220
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 282 SMYNDPOF-----GTSCEITGFGKENSTDYLPOLKMTVVVKLISHRECQPHYYGSEVTT 337
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 221 S-----PNYVVADKTVCVITGWGETGT--FGVGRLEKARLPVIENKVCNRYEYLNGRVKS 274
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 338 KMLCADPQWKTDSCGDSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVTVVSHFLP 397
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 TELCAGLAGTSCGDSGGPLVCFEKKYILQGVTSWGLGCARPNKPGVYVRVSTVVP 334
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 398 WI 399
  ||
Db 335 WI 336

RESULT 27
ID PLMN BOVIN STANDARD; PRT; 812 AA.
AC P06868; Q28162;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7).
GN PLG.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Liver;
RA Berglund L., Andersen M.D., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen cDNA.";
RL Int. Dairy J. 5:593-603(1995).
RN [2]
RN SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.
RP MEDLINE=85203906; PubMed=3846532;
RX Schaller J., Moser P.W., Danegger-Muller G.A.K., Rosselet S.J.,
RA Kampfer U., Rickli E.E.;
RT "Complete amino acid sequence of bovine plasminogen. Comparison with
RT human plasminogen.";
RL Eur. J. Biochem. 149:267-278(1985).
RN [3]
RN SEQUENCE OF 706-812 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [4]

```

RP CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=88185329; PubMed=3356193;  
RA Marti T., Schaller J., Rickli B.E., Schmid K., Kamerling J.P.,  
RA Gerwig G.J., van Halbeek H., Vliegthart J.F.;  
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine  
RT plasminogen. Species specificity in relation to sialylation and  
RT fucosylation patterns.";  
RL Eur. J. Biochem. 173:57-63(1988).  
CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS  
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING  
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION  
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE  
CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN  
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH  
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,  
CC LAMININ AND VON WILLEBRAND FACTOR.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;  
CC higher selectivity than trypsin. Converts fibrin into soluble  
CC products.  
CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN  
CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO  
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.  
CC -!- PTM: N-LINKED GLYCANS CONTAIN N-ACETYLGLUCOSAMINE AND SIALIC ACID.  
CC O-LINKED GLYCANS CONSIST OF GAL-GALNAc DISACCHARIDE WITH IS  
CC MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).  
CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN  
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.  
CC -!- SIMILARITY: Contains 5 kringle domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL; X79402; CAA55939.1; -;  
DR EMBL; K02935; AAA30714.1; -;  
DR PIR; S45046; PLBEO.  
DR HSSP; P00747; 2PK4.  
DR MEROPS; S01.233; -;  
DR GlycoSuiteDB; P06868; -;  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR InterPro; IPR003966; Prothrombin.  
DR Pfam; PF0001254; Ser\_protease\_Try.  
DR Pfam; PF000051; kringle; 5.  
DR Pfam; PF00024; PAN; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR01505; PROTHROMBIN.  
DR ProDom; PD000395; Kringle; 5.  
DR SMART; SM00130; KR; 5.  
DR SMART; SM00473; PAN AP; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 5.  
DR PROSITE; PS00070; KRINGLE\_2; 5.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;  
KW Signal.  
FT SIGNAL. 1 26  
FT CHAIN 27 812 PLASMINOGEN.  
FT CHAIN 27 583 PLASMIN HEAVY CHAIN A.  
FT CHAIN 584 812 PLASMIN LIGHT CHAIN B.  
FT DOMAIN 110 188 KRINGLE 1.

FT	DOMAIN	192	269		KRINGLE 2.
FT	DOMAIN	282	359		KRINGLE 3.
FT	DOMAIN	384	461		KRINGLE 4.
FT	DOMAIN	584	564		KRINGLE 5.
FT	DOMAIN	584	812		SERINE PROTEASE.
FT	CARBOHYD	315	315		N-LINKED (GLCNAC. . .).
FT					/FTID=CAR_000014.
FT	CARBOHYD	365	365		O-LINKED (GALNAC. . .).
FT					/FTID=CAR_000015.
FT	ACT_SITE	624	624		CHARGE RELAY SYSTEM.
FT	ACT_SITE	667	667		CHARGE RELAY SYSTEM.
FT	ACT_SITE	762	762		CHARGE RELAY SYSTEM.
FT	CONFLICT	335	335		N -> D (IN REF. 2).
FT	CONFLICT	516	516		Q -> H (IN REF. 2).
FT	CONFLICT	555	555		P -> L (IN REF. 2).
FT	CONFLICT	744	744		T -> R (IN REF. 3).
SQ	SEQUENCE	812 AA;	91216 MW;	386AA691E220946	CRC64;

Query Match 21.1%; Score 484.5; DB 1; Length 812;  
Best Local Similarity 34.9%; Pred. No. 9e-31;  
Matches 124; Conservative 41; Mismatches 151; Indels 39; Gaps 11;

QY	50	CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHYCRNPD-NR	108
Db	485	CMIGTGKSYRGKXATTVAGVPCQEWAAQEPHQHSIFTPETNP-QSGL-ERNYCRNPDGV	542
QY	109	RRPCYQVQGLKPLVQECMHVDCADGKKPSPPEELKFCQCKTLRPR---FKIIGSEFT	165
Db	543	NGPCYNTMPKPP-DYCDVPQC-----ESSFDCGKPKVPEKKCSGRIVGCVS	590
QY	166	TENQPFAAIYRRHRGGSVYVCGSLISPCWVISATHCIFDYPKKEDYIVYGRSLN	225
Db	591	KPHSWPQVSLRRSSR-----HFCGGLTISPQWLTAAHCLDNLALSFKYVILGAHNEK	645
QY	226	SNTQGEKMFVENILHKQVYADTLAHNDIALKIRSKGRCQAPQRTTQTCICLPSMYN	285
Db	646	VREQSVQEIPIVSRLEFRPSQA-----DIALKL---SRPAITKEVIPACLPFPNY	693
QY	286	DPQFGTSCETIGFGKENSTLYPE-OLKMTVVKLISHRECOQPHYGVSEVTTKWLCAAD	344
Db	694	MVAARTECIYTGWGETQGT---FOEGLLKAHLVNIENKVCNRNEYLDGRVKPTELCAH	750
QY	345	PQWKTDSCQSDSGGGLVCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSVHFLPMI	399
Db	751	LIGGTDSCQSDSGGGLVCFEKDKYILQGVTSWGLGCARPKNKPGVYVRSVPMI	805

RESULT 28  
APOA MACMU  
ID APOA MACMU STANDARD; PRT; 1420 AA.  
AC P14417;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).  
GN LPA.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_taxid=9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89174660; PubMed=2925643;  
RA Tomlinson J.E., McLean J.W., Lawn R.M.;  
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of  
RT synthesis.";  
RL J. Biol. Chem. 264:5957-5965(1989).  
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)  
CC (Lp(a)). It has serine proteinase activity and is able of  
CC autolysis. Inhibits tissue-type plasminogen activator 1.  
CC Lp(a) may be a ligand for megalin/Gp 330.  
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and





```

RESULT 30
NETR HUMAN
AC P56730; Q9UP16; STANDARD; PRT; 875 AA.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurotysin precursor (EC 3.4.21.-) (Motopsin) (Leydin).
GN PRSS12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98201705; PubMed=9540828;
RA Proba K., Gschwend T.P., Sonderegger P.;
RT "Cloning and sequencing of the cDNA encoding human neurotysin.";
RL Biochim. Biophys. Acta 1396:143-147(1998).
RN [2]
SEQUENCE OF 615-875 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=99203523; PubMed=10103056;
RA Poorafshar M., Hellman L.;
RT "Cloning and structural analysis of leydin, a novel human serine
protease expressed by the Leydig cells of the testis.";
RL Eur. J. Biochem. 261:244-250(1999).
CC -!- FUNCTION: PLAYS A ROLE IN NEURONAL PLASTICITY AND THE PROTEOLYTIC
ACTION MAY SUBSERVE STRUCTURAL REORGANIZATIONS ASSOCIATED WITH
LEARNING AND MEMORY OPERATIONS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Brain and Leydig cells of the testis.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 4 SRCR domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; AJ701531; GAA04816.1; -;
DR EMBL; AR077298; AAD25919.1; -;
DR HSSP; P00763; IDPO.
DR Genew; HGNC:9477; PRSS12.
DR MIM; 606709; -;
DR MEROPS; S01.237; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00530; SRCR; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00258; SPERACTRCPTR.
DR Prodom; PD000395; kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00202; SR; 4.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00707; KRINGLE_2; 1.
DR PROSITE; PS00420; SRCR_1; 3.
DR PROSITE; PS00287; SRCR_2; 4.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.

```

FT	SIGNAL	1	20	POTENTIAL.
FT CHAIN	21	875	NEUROTRYPSIN.	
FT DOMAIN	23	92	PRO-RICH.	
FT DOMAIN	93	165	KRINGLE.	
FT DOMAIN	170	271	SRCR 1.	
FT DOMAIN	280	381	SRCR 2.	
FT DOMAIN	387	487	SRCR 3.	
FT DOMAIN	500	601	SRCR 4.	
FT DOMAIN	619	875	SERINE PROTEASE.	
FT DOMAIN	619	875	ZYMOGEN ACTIVATION REGION.	
FT ACT SITE	630	631	REACTIVE BOND (POTENTIAL).	
FT ACT SITE	676	676	CHARGE RELAY SYSTEM.	
FT ACT SITE	726	726	CHARGE RELAY SYSTEM.	
FT ACT SITE	825	825	CHARGE RELAY SYSTEM.	
FT DISULFID	619	750	POTENTIAL.	
FT CARBOHYD	26	26	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD	683	683	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CONFLICT	663	663	A -> V (IN REF. 2).	
FT CONFLICT	701	701	E -> V (IN REF. 2).	
FT CONFLICT	839	841	VVY -> AAL (IN REF. 2).	
SQ SEQUENCE	875 AA; 97011 MW; B66EC946DC208DC8 CRC64;			

Query Match 20.1%; Score 462; DB 1; Length 875;  
Best Local Similarity 29.8%; Pred. No. 5.9e-29;  
Matches 130; Conservative 60; Mismatches 156; Indels 90; Gaps 16;

QY	33	CPKFGGQH-C--EIDKSKTCYEG-----	-----NGHF----	57
DB	465	CSRQWGRHDCSHREDVSIACYPGGEGHRLSLGFPVRLMDGKKEGRVEVFINGQGTI	-----	524
QY	58	-----YRGKASTDTM-----GRPCLPMNSATVLOQTYHAHRSDALQGL	-----	96
DB	525	CDDGWTDKDAVICQLGYKGFARATMAYFEGKGPPIHVDNVKCTGNERSLADCIQDI	-----	584
QY	97	GKHNYCRNPDRRRPWCYVQVGLKPLVQECMVHDCADGKKPS--SPPEELKFCQGOKTLR	-----	154
DB	585	GRHN-CRHSEDAGVICDYF-----GKKASGNSNKESLSVVCGLRLH	-----	625
QY	155	PRFK-IIGGEFTIENQPFPAIYRRHROGSVTVYCGSLISPCWVISATHCFIDYPKK-	-----	212
DB	626	RQKRIIGKNSLRGGWPQVSLRLKSHDGLRLLCGATLLSSCWLTAAHCFKRYGNST	-----	685
QY	213	EDYIVVLRSLNSNTQGMKFEVENLIILHKDYSADTLAHNDIALLKIRSKRGCAQPS	-----	272
DB	686	RYAVRVGYHTLVPEEFEEIGVQOIVIHREYRDRSDY--DIALVRLQGGPEEQCARFS	-----	743
QY	273	RTIQITCLPSMYNDPQ-FQTSCEITFGKENSTDLYPEQLKMTVVKLISHRECCQPHY	-----	331
DB	744	SHVLPACLPLWRERPOKTASNCYITGWG--DTGRAYSRTLQAAIPLLPKRFCEE--RY	-----	798
QY	332	GSEVTTKMLCAAD--PQWKTDSQGDSCGGLVCSLQGR-MTLTGIVSWGRGCALKDKPGV	-----	388
DB	799	KGRFTGRMLCAGNLHEHKRVDSQGDSCGGLPLMCCERPGEWSVWYVYVTSWGYCGV	-----	858
QY	389	YTRVSHFLPWIRSHTK 404		
DB	859	YTKVSFAFVPIKSVTK 874		

Search completed: December 3, 2003, 14:40:06  
Job time : 15.2061 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:23 ; Search time 50.5697 Seconds  
(without alignments)  
2097.294 Million cell updates/sec

Title: US-09-880-503-3  
Perfect score: 2301  
Sequence: 1 SNEHQVPSNCDLNGTCV.....VSHFLPWIRSHYKENGAL 411

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:\*  
1: sp archaea:\*  
2: sp bacteria:\*  
3: sp fungi:\*  
4: sp human:\*  
5: sp\_invertebrate:\*  
6: sp mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	1923	83.6	433	Q8MIL0	Q8milo11 cryptotolagus
2	1918	83.4	433	Q8MIL0	Q8milo11 cryptotolagus
3	982.5	42.7	214	Q8MIL0	Q8milo11 cryptotolagus
4	858.5	37.3	516	Q8MIL0	Q8milo11 cryptotolagus
5	842.5	36.6	231	Q8MIL0	Q8milo11 cryptotolagus
6	826.5	35.9	559	Q8MIL0	Q8milo11 cryptotolagus
7	819	35.6	564	Q8MIL0	Q8milo11 cryptotolagus
8	816.5	35.5	562	Q8MIL0	Q8milo11 cryptotolagus
9	784.5	34.1	395	Q8MIL0	Q8milo11 cryptotolagus
10	778	33.8	154	Q8MIL0	Q8milo11 cryptotolagus
11	726	31.6	653	Q8MIL0	Q8milo11 cryptotolagus
12	693.5	30.1	616	Q8MIL0	Q8milo11 cryptotolagus
13	689	29.9	615	Q8MIL0	Q8milo11 cryptotolagus
14	677	29.4	597	Q8MIL0	Q8milo11 cryptotolagus
15	677	29.4	597	Q8MIL0	Q8milo11 cryptotolagus
16	674.5	29.3	517	Q8MIL0	Q8milo11 cryptotolagus

17	661.5	28.7	560	4	Q14520	Q14520 homo sapien
18	585	25.4	128	6	Q97587	Q97587 cryptotolagus
19	501	21.8	501	11	Q91WJ5	Q91WJ5 mus musculus
20	498	21.6	810	4	Q15146	Q15146 homo sapien
21	496	21.6	454	6	Q46506	Q46506 papio hamad
22	496	21.6	812	11	Q9R0W3	Q9R0W3 rattus norv
23	495.5	21.5	300	4	Q96EF3	Q96EF3 homo sapien
24	495	21.5	103	6	Q95M89	Q95M89 equus caball
25	491.5	21.4	429	13	Q8AVB0	Q8AVB0 brachydanio
26	482.5	21.0	334	6	Q46507	Q46507 papio hamad
27	472.5	20.5	806	6	Q18783	Q18783 macropus eu
28	463.5	20.1	868	5	Q9Y1V3	Q9Y1V3 polyandroca
29	442.5	19.2	327	4	Q8N171	Q8N171 homo sapien
30	435.5	18.9	284	4	Q8NF86	Q8NF86 homo sapien
31	433.5	18.8	761	11	Q99JCB	Q99JCB rattus norv
32	426.5	18.5	267	5	Q9BK47	Q9BK47 luidia foli
33	425.5	18.5	505	5	Q966V4	Q966V4 halocynthia
34	424	18.4	537	4	Q9BYE1	Q9BYE1 homo sapien
35	420	18.3	471	11	Q8CFE0	Q8CFE0 mus musculu
36	420	18.3	581	4	Q9BYE2	Q9BYE2 homo sapien
37	420	18.3	802	4	Q8IUE2	Q8IUE2 homo sapien
38	420	18.3	811	4	Q8IU80	Q8IU80 homo sapien
39	419.5	18.2	799	11	Q8DBI0	Q8DBI0 mus musculu
40	414	18.0	638	11	Q8R0P5	Q8R0P5 mus musculu
41	411.5	17.9	276	11	Q8CGR6	Q8CGR6 mus musculu
42	409	17.8	371	11	Q8CJL6	Q8CJL6 rattus norv
43	409	17.8	445	11	Q8CJL7	Q8CJL7 rattus norv
44	405.5	17.6	422	4	Q8WVC1	Q8WVC1 homo sapien
45	402	17.5	455	11	Q8CDR0	Q8CDR0 mus musculu

## ALIGNMENTS

### RESULT 1

Q8MIL0  
ID Q8MIL0 PRELIMINARY; PRT; 433 AA.  
AC Q8MIL0;  
DT 01-OCT-2002 (TREMELrel. 22, Created)  
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)  
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
DE Urokinase-type plasminogen activator.  
GN PLAU  
OS Cryptolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22155945; PubMed=12149463;  
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,  
RA Dichek D.A.;  
RT "Increased expression of urokinase during atherosclerotic lesion  
RT development causes arterial constriction and lumen loss, and  
RT accelerates lesion growth."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; AY122285; AAM83187.1; -!  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Ser protease\_Try.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.

```

DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65E64F36415549B0 CRC64;

Query Match      83.6%; Score 1923; DB 6; Length 433;
Best Local Similarity 83.8%; Pred. No. 5.3e-177;
Matches 346; Conservative 25; Mismatches 40; Indels 2; Gaps 1;

QY 1 SNELHGV--PSNCDCLNGGTCVSNKYFSNIHWCNCPKKGQHCETDKSKTCYEGNGHFY 58
DB 21 SHELHGVSDASNGCGLNGGTCVTYKYFSNIWRCNCPKKGQHCETDITLTCYHGDGHSY 80

QY 59 RKASTDITMGRPCLPWNSATVLQOITYHAHRSALQLGLGKHNYCRNPDNRRCWCVQVG 118
DB 81 RGRANTDINDRPLCLAWNSANVLTKTYHAHRPDALQLGLGKHNYCRNPDHQRRCWCVQVG 140

QY 119 LKPLVQECVHDCADGKSPPEELKFCQGOKTLRPRKLIIGGEFTTIENQPFAAIYR 178
DB 141 LKQLIQECKVHDCSSGKKPALPPGKLEFCQGOKALRPRFKIIGGEFTTIENQPFAAIYR 200

QY 179 RHGGSVTVYVCGSLSPCWVVSATHCFIDYPKKEDYIVVLGRSLNSNTQGMKFEVEQ 238
DB 201 RHGGSVTVYVCGSLSPCWVVSATHCFINHQKKEDYIVVLGRSLNSMTPGEMKFEVEQ 260

QY 239 LILHKDYSADTLAHNDIALKLRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGF 298
DB 261 LILHEGYSADTLAHNDIALKILSNNGCAQPSRSIQTICLPWNADPNFGTSCEITGF 320

QY 299 GRENSTDYLYPEQLKMTVVKLVSHRECCQPHYYSVETTKMLCAADPQWETDSCQDSSG 358
DB 321 GRENSTDYLYPEQLKMTVVKLVSYQCQPHYYSVETTKMLCAADPQWETDSCQDSSG 380

RESULT 2
Q8MHY7 PRELIMINARY; PRT; 433 AA.
AC Q8MHY7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Yano W.; Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA, complete cds."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AV029517; AAA40239.1; -.
DR EMBL; AB087224; BAC02685.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.

```

```

DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 5DD35A371010A5EE CRC64;

Query Match      83.4%; Score 1918; DB 6; Length 433;
Best Local Similarity 83.5%; Pred. No. 1.6e-176;
Matches 345; Conservative 25; Mismatches 41; Indels 2; Gaps 1;

QY 1 SNELHGV--PSNCDCLNGGTCVSNKYFSNIHWCNCPKKGQHCETDKSKTCYEGNGHFY 58
DB 21 SHELHGVSDASNGCGLNGGTCVTYKYFSNIWRCNCPKKGQHCETDITLTCYHGDGHSY 80

QY 59 RKASTDITMGRPCLPWNSATVLQOITYHAHRSALQLGLGKHNYCRNPDNRRCWCVQVG 118
DB 81 RGRANTDINDRPLCLAWNSANVLTKTYHAHRPDALQLGLGKHNYCRNPDHQRRCWCVQVG 140

QY 119 LKPLVQECVHDCADGKSPPEELKFCQGOKTLRPRKLIIGGEFTTIENQPFAAIYR 178
DB 141 LKQLIQECKVHDCSSGKKPALPPGKLEFCQGOKALRPRFKIIGGEFTTIENQPFAAIYR 200

QY 179 RHGGSVTVYVCGSLSPCWVVSATHCFIDYPKKEDYIVVLGRSLNSNTQGMKFEVEQ 238
DB 201 RHGGSVTVYVCGSLSPCWVVSATHCFINHQKKEDYIVVLGRSLNSMTPGEMKFEVEQ 260

QY 239 LILHKDYSADTLAHNDIALKLRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGF 298
DB 261 LILHEGYSADTLAHNDIALKILSNNGCAQPSRSIQTICLPWNADPNFGTSCEITGF 320

QY 299 GRENSTDYLYPEQLKMTVVKLVSHRECCQPHYYSVETTKMLCAADPQWETDSCQDSSG 358
DB 321 GRENSTDYLYPEQLKMTVVKLVSYQCQPHYYSVETTKMLCAADPQWETDSCQDSSG 380

QY 359 PLVCSLQGRMTLTGIVSWRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
DB 381 PLVCSVQGRMTLTGIVSWRGCAKDKPGVYTRVSRFLPWIRSHTKKEENGLAL 433

RESULT 3
Q9XT70 PRELIMINARY; PRT; 214 AA.
ID Q9XT70;
AC Q9XT70;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Yano W.; Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA, complete cds."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AV029517; AAA40239.1; -.
DR EMBL; AB087224; BAC02685.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.

```





```
Query Match      36.6%; Score 842.5; DB 11; Length 231;
Best Local Similarity 71.5%; Pred. No. 4.7e-73;
Matches 143; Conservative 23; Mismatches 33; Indels 1; Gaps 1;

QY 9 SNCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFGKASTDTMG 68
Db 30 SNGCGQNGGVCSYKVFSTRRCSPKXFOGEHCEIDASKTCYHGNGDSYRGKANTDTKG 89

QY 69 RPCLPNSATVLQOQTYHAHRSALQLGLGKHNYCRNPDNRNRRPWCYVQGLKPLVQECMV 128
Db 90 RPLAWNAFVLOKPNYNAHSPDAISLGLGKHNYCRNPDNQKRPWCYVQIGLRQVQECMV 149

QY 129 HDCADGKPPSPPEELKFCQGOQKTLRPRFKIIGSEFTTIENQWFAALYRHRGGSVTY 187
Db 150 HDCSLSKSSVSDQFQCGQKALRPRFKIVGSEFTTEVENQWFAALYKNNKGGSPPSF 209

QY 188 VCGSLISPCWVISAHCIFI 207
Db 210 KCGSLISPCWASAAHCIFI 229

RESULT 6
Q91VP2 PRELIMINARY; PRT; 559 AA.
AC Q91VP2;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
GN PLAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC011256; AAH11256.1; -.
DR HSSP; P00761; IANI.
DR MGD; MGI:97610; Plat.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD00015; KRINGLE.
DR PRODOM; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRINCTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
SQ Serine protease.
SEQUENCE 559 AA; 63122 MW; 8CCE2BDB94514D9 CRC64;

Query Match      35.9%; Score 826.5; DB 11; Length 559;
Best Local Similarity 35.7%; Pred. No. 5.1e-71;
Matches 176; Conservative 65; Mismatches 155; Indels 97; Gaps 10;

QY 3 ELHVP-----SNCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFGY 58
Db 74 QCHSVPRSCSEPRCNGGTCQALIFSDP-VQCQPDGFVGRKCDIDTRATCFEEQGIT 132

QY 59 RGKASTDTMGRPCLPNWSATVLQOQTYHAHRSALQLGLGKHNYCRNPDNRNRRPWCYVQVG 118
Db 133 RGTMTAESAEGACINWNSVLSUKPYNARRPNAIKLGLGNHNYCRNPDRLKPCWYVFKA 192

QY 119 LKPLVQECMVHDCADGKPS----- 138
Db 193 GKYTFEFTCPACPKGSEDCYVKGVTYRGTHSLTTSQASCLPWNISVLIMGKSYTAWRT 252

QY 139 -----SPPEELK-----FQCG-OKTLRPRFKIIGG 162
Db 253 NSQALGLGRHNYCRNPDGDARPWCHVMKDKLWECYDMSPCSTCGLRQYKRFQRIKGG 312

QY 163 EFTTIENQWFAALY-RRHRGGSVTYVCGSLISPCWVISAHCIFIDYPKKEDYIVYLGR 221
Db 313 LYTDITSHPWQAPIFVKNKRSPGERELCGVLISSCWLSAAHCFLEPRPPNHLKVVLGR 372

QY 222 SRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLLKIRSEKGRCAQPSRTIQTICLP 281
Db 373 TYRVVPEBEEQTEIEKYIVHEEFDDET--YNDIALQLRSQSKCAQESSVGTACL 430

QY 282 SMYNDPQF----GTSCEITGFGKENSTDYLYPQLKMTVVKLISHRECCQPHYVGSEVTT 337
Db 431 ----DNLQLPDWTECELSGKGHEASSPFFSDRLKEAHVRLYPSSRCTSOHLFNKVTIN 486

QY 338 KMLCAADP-----QWKTDCQDGGGGLPVCSLQGRMTLTGIVSWGRCALKKRPGYVTRV 392
Db 487 NMLCAGDTSGGNQDLHDACQDGGGLVCMINKQMTLTGIIISWGLGCGKQDVPGVYTKV 546

QY 393 SHFLPWIRSHRSTKE 405
Db 547 TNYLDWIHDNMKQ 559

RESULT 7
Q8MKB1 PRELIMINARY; PRT; 564 AA.
AC Q8MKB1;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Tissue-type plasminogen activator.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M, Yoshida E, Anai K, Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AY029518; AAK40240.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD00015; KRINGLE.
DR PRODOM; PD000395; Kringle; 2.
```

```

DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 564 AA; 62726 MW; 459D8BAC6DA937C CRC64;

Query Match      35.6%; Score 819; DB 6; Length 564;
Best Local Similarity 35.7%; Pred. No. 2.8e-70;
Matches 178; Conservative 55; Mismatches 148; Indels 118; Gaps 9;

QY 5 HOVP-SNCD---CLNGGTCVSNKYFNIHWCNPKKFGGQHCEID-----45
Db 80 HSPVQSCSEPRCLNGTCSQALYFSDP-VCQPEGFVGRCEVDTRACYEDRGIGYRG 138
QY 46 -----45
Db 139 TWSSTESGAQCVNNSWLALPKYPGCRKPNALRLGLGNHNYCRNPDRDTKPNCYVFRAGT 198
QY 46 -----KSKTCYEGNGHYRGKASTDTMGRLPWNSTVTLQOYTHAHRSDA 91
Db 199 YSPFCSTPACSKENGNCGYLKGQAYRGTHSLTTSGLSCPLPWNMMLLVGEKYTEARQNSA 258
QY 92 LQGLGKHNYCRNPONRRPWCVVQGLKPLVQECMVHDCADGKXPSPPPEELKFCQG-Q 150
Db 259 EALGLGKHNYCRNPDGSKPWCHLVKNRLVYCDVPCA-----TCGR 304
QY 151 KTLRPRFKIIGEFTTIENQPFPAIY-RRHGGSVTYVCGSLISPCWVISATHCFIDY 209
Db 305 QDKQPFRIKGLGFTDITAHPPQAAIFVNNRSPGRLFCGGLINSCWVLSAAHCFLE 364
QY 210 PKEDYIVVGLGRSLNSNTQGMKEVENLILHKDYSADTLAHNDIALLKIRSKRGCA 269
Db 365 FPOQKRLVILGRITYPLVSAEEQIPEVEQPIILHERFDEGT--YNDIALLLKSTSGSCA 422
QY 270 QSRITQITCLPSMYNDPQFGTSCITGFGKENSTDYLPVOLKMTVVKLIASHRCQQPH 329
Db 423 QESQAVRLVCLPDASLQLPDWTCELSGKGKHEEFPVSEQLKAHVLYPSSRCTPOQ 482
QY 330 YGSEVTTNMLCAADPQW-----KTDSQCGSGGGLVCSLQGRMTLTGIVSWGRCALK 383
Db 483 LKNRTVTGNMLCAGDTRSGGAQVNLHDAQCGDGGGLVCMTDGHTLIGIISWGLGCGQK 542
QY 384 DKPGVYTRVSHPLPWIRSH 402
Db 543 DVPGVYTKVNYLGIQHH 561

RESULT 8
Q8SQ23 PRELIMINARY; PRT; 562 AA.
AC Q8SQ23;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Enamel organ;
RA Ding Y., Xue J., Bartlett J.D.;

```

RT "T-plasminogen activator in tooth tissues.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.

DR EMBL; AF364605; AAM00297.1; -.
DR HSSP; P00761; 1AN1.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser\_protease\_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp\_Spc; 1.
DR PROSITE; PS00022; EGF\_1; 1.
DR PROSITE; PS01186; EGF\_2; 1.
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.
DR PROSITE; PS00021; KRINGLE\_1; 2.
DR PROSITE; PS00070; KRINGLE\_2; 2.
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.
DR PROSITE; PS00135; TRYPSIN\_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 562 AA; 63668 MW; F9B6B4C7CB101E8 CRC64;

Query Match 35.5%; Score 816.5; DB 6; Length 562;
Best Local Similarity 35.2%; Pred. No. 4.8e-70;
Matches 172; Conservative 61; Mismatches 166; Indels 89; Gaps 8;

QY 3 ELHQP-SNCD---CLNGGTCVSNKYFNIHWCNPKKFGGQHCEIDKSKTCYEGNGHY 58
Db 77 QCHSVFVSCSPRCFNGGTCLQALYFSDP-VCQCPVGFQIGQCEIDARATCYEDQGIT 135
QY 59 RKASGTDITMGRLPWNSTVTLQOYTHAHRSDALQGLGKHNYCRNPDRRRPWCVVQVG 118
Db 136 RGTWSTTESGAFCVNVNWTSGLASMPYNGRPPDAVLGLGNHNYCRNPDKDSPWCYIFKA 195
QY 119 LKPLVQECMVHDCADGKXPSPPPEELKFO-----147
Db 196 EKSPDFCSTPACTKEKECEYTGKGLDYRGTRSLTWSGAFCLPWISLVLMGKIYTAWNSN 255
QY 148 -----CGQKTLR-PRFKIIGGE 163
Db 256 AQTGLGKHNYCRNPDPGDTQPWCHVLKDHKLTWKEYCDLPQCVCGLRQYKEPQFRIKGL 315
QY 164 FTTIENQPFPAIY-RRHGGSVTYVCGSLISPCWVISATHCFIDYPKKDYIVVGLRS 222
Db 316 YADITSHPWQAAIFVNNRSPGRLFCGGLISSCWVLSAAHCFQERPPPHVVRVVLGT 375
QY 223 RLNSNTQGMKEVENLILHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTCLPS 282
Db 376 YRLVPEEEQAFVEKEYIVHKEFFDDDT--YNDIALQLKSDSLTCAQESDAVRTVCLPE 433
QY 283 MYNDPQFGTSCITGFGKENSTDYLPVOLKMTVVKLIASHRCQQPHYVGSSEVTTNMLCA 342
Db 434 ANLQLPDWTECELSGKGKHEASSFPYSERLKAHVRLYPSSRCTSKHLFNKTIITNNMLCA 493
QY 343 ADPQW-----KTDSQCGSGGGLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 396
Db 494 GDRSGGDNANLHDACQDGGGLVCMKGNHMTLVGVISWGLGCGQKDPVGYTKVNYL 553
QY 397 PWIRSHTK 404
Db 554 NWIRDNTR 561

RESULT 9



```
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -.
DR HSP; P00761; IAN1.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase;
KW Serine protease.
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 31.6%; Score 726; DB 11; Length 653;
Best Local Similarity 37.3%; Pred. No. 3.2e-61;
Matches 158; Conservative 57; Mismatches 167; Indels 42; Gaps 9;

QY 5 HQVPSNCDLNGGTCVSNKYFSNIHW-----CWCPRKFGQHQHCEIDKSKTCYEGNGH 56
DB 239 HTACLSSPCLNGGTC-----HLVGTGTSVCTCPYAGRCFNIVPTBHCFLNGT 289
QY 57 FYRGKASTDTWGRPCLPNWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRERREWCVYQ 116
DB 290 EYRGVASTAASGLSCLANWSLLYQELHVDVSAAAVLLGLGPHAYCRNPDNDRPWCYV 349
QY 117 VGLKPLVCEWVHDCAD-GKRPSSPEEL-----KPCGQK-----TLRPRFII 160
DB 350 KDNALSWCYRLTACESLARVHSQSFETLALPESAPAVRPTCGRRHKRTFLRPR--II 407
QY 161 GGEFTTIENQPMFAAIYRRHRGGSVTVYCGSLISFCVWISATHCFIDYPKKEDIVVLG 220
DB 408 GGSLSLPGSHFWLAIIY---IGNS---FCAGSLVHTCWVNSAHFANSPRDSITVVLG 461
QY 221 RSLRNSNTQGMKFEVFNILHKYSDATLHAHNDIALLKIRSKBGRCAQPSRTIQTICL 280
DB 462 QHFFNRRTDVTQTFGIEKYVPVTLYSVENNNH-DLVLRLLKKGECRCAVRSQFQPICL 520
QY 281 PSMYNDPQFGSCBTGFGKENSIDYLPQELKMTVVKLISHRECOQPHYGSEVTTTML 340
DB 521 PEAGSSFTGHKQCIAGHGHMDENVSSYNSLLEALVPLVADHKCSSPEVYGADISPNWL 580

341 CAADPQWKTDSCGDSGGPLVCSLQGRMTLTGIVSWGRGALKDKFGVTVTVSHFLPWIR 400
DB 581 CAGYFDCKSDACQDSGGPLVCEKNGVAYLYGIISWGDGCGRLNKFGVTVTVRANYVDWIN 640
QY 401 SHTK 404
DB 641 DRIR 644

RESULT 12
O97507 PRELIMINARY; PRT; 616 AA.
AC O97507;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE FXII.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Takahashi T., Kihara T.;
RT "Porcine liver factor XII.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AB022426; BAA37148.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.211; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBD1A9 CRC64;

Query Match 30.1%; Score 693.5; DB 6; Length 616;
Best Local Similarity 35.3%; Pred. No. 4e-58;
Matches 161; Conservative 61; Mismatches 159; Indels 75; Gaps 14;

QY 6 QVPNSCDLNGGTCVSNKYFSNIHWNCPCPKFGQHQHCEIDKSKTCYEGNGHFGKASTD 65
```

```
Db 176 QVCTNPCLNGSGCLOTE---GHRCLRCPTGYAGRLCDVLDKERCYSRGLSYRGMAQTT 232
QY 66 TMGRPCLPWN SATVLOQTY-HAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLKPLVQ 124
Db 233 LSGAPCPWAS---EATYWNMTAEQALNWLGDHAFCRPNDDNTRWCFWFGDQLSWQ 288
QY 125 ECMVHDCAD-GKKP-----SSPPEEL-144
Db 289 YCLARCAQPIGEAPILLTPTQSPSEHDSPLLSREPOPTTQPSQNLTSAMCAPPBORG 348
QY 145 -----KFOGQKTLRPF---KIIGBEFTTIENQWFAAIYRRHGGSVTVVCGSGLI 194
Db 349 PLPSAGLVGCGQR-LRRLSSLNRIVGGLVALPOAHPIYIAALYWGQ-----FCAGSLI 401
QY 195 SPCWVISATCFIDYPKEDYIVYGRSRLNSNTQSGEMKFEVENLILHKDYSADTLAHNN 254
Db 402 APCWLVTAACHLQNRPAPEELTVVLGQDRHNQSCQCTLAVERSYRLHESYSPKTYQH-- 459
QY 255 DIALLKTR-SKEGRCAQPSRTIQTICLP---SMYNDQFGTSCITGFGKENSNDYLYPE 310
Db 460 DLALVRLKETADGCAHSPFPVPCVCLPRSVASSAEPE-GALCEVAGWGHQFEGAEYSS 518
QY 311 QLKMTVVKLISHRECQPHYVGSVTTKMLCAADPOWKTDSCQDGGPLVC---SLQGR 367
Db 519 FLQEAQVPLISPERCSAADVHGAAFTPGMLCAGFLEGGTDACQDGGPLVCEDETAERQ 578
QY 368 MTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHT 403
Db 579 LVLRGIVSWGCGDRCLKPGVYTDVANYLAWIQEHT 614

RESULT 13
Q81ZZ5 PRELIMINARY; PRT; 615 AA.
AC Q81ZZ5;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Coagulation factor XII-Mie.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wada H., Nishioka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;
RT "Molecular characterization of coagulation factor XII-Mie.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095845; BAC23095.1; -
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C72DB CRC64;

Query Match 29.9%; Score 689; DB 4; Length 615;
Best Local Similarity 34.3%; Pred. No. 1.1e-57;
Matches 153; Conservative 59; Mismatches 164; Indels 70; Gaps 9;

QY 13 CLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db 183 CLHGGRCLE---VEGHRCLCHCPVGTGPFCDVDTKASCYDGRGLSYRGLARTTLSGAPCQ 239
QY 73 PWN SATVLOQTY-HAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMVHDC 131
Db 240 PWAS---EATYENVTAEQARNWGLGHAFCRPNDDNIRPWCFLNDRLSWEYCDLAQC 295
QY 132 -----ADGKFPSSPEEL 144
Db 296 QTPTQAPPTVPSPRLHVLMPAQAAPPKPPQPTTRTPQSQTPGALPAKREQPSLTRNG 355
QY 145 KFOGO---KTLRPRFKIIGBEFTTIENQWFAAIYRRHGGSVTVVCGSGLISPCWVIS 201
Db 356 PLSCGQRLKSLSMTRKRVVGLVALRGHAHPYIAALYWGHS-----FCAGSLIAPCWILT 409
QY 202 ATHCFIDYPKEDYIVYGRSRLNSNTQSGEMKFEVENLILHKDYSADTLAHNDIALKI 261
```

```
Db 410 AAHCLQDRPAPEDLTVLVGQERRNHSCEPCQTLAVRSYRLHEAFS---PVSQYHDLALLRL 467
QY 262 R-SKEGRCAQPSRTIQTICLPSPMYNDQFGTSCITGFGKENSNDYLYPEOLKMTVVVKLI 320
Db 468 QEDADSCALLSPYQVPCVLPSSGAARPSETTLQVACGQHGFEGAEYASFLQEAQVFFL 527
QY 321 SHRECQPHYVGSVTTKMLCAADPOWKTDSCQDGGPLVCSLQG---RMTLTGIVSWG 377
Db 528 SLERCAPDVHGSSILPGMLCAGFLEGGTDACQDGGPLVCEPQAAERLTLQGIISWG 587
QY 378 RGALXDKDPQVYTRVSHFLPWIRSHT 403
Db 588 SGGCDRKNKPGVYTDVAYLAWIREHT 613

RESULT 14
Q9TVAB PRELIMINARY; PRT; 157 AA.
AC Q9TVAB;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21071388; PubMed=11204721;
RA Balcerzak D., Querengesser L., Dixon W.T., Baracos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
activators and inhibitors in bovine skeletal muscle.";
RL J. Anim. Sci. 79:94-107(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF1474761; AAD30301.1; -
DR HSP; P00749; IURK.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle_1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kinase; Kringle.
FT NON_TER 1
FT NON_TER 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 29.4%; Score 677; DB 6; Length 157;
Best Local Similarity 73.9%; Pred. No. 2.7e-57;
Matches 116; Conservative 15; Mismatches 26; Indels 0; Gaps 0;

QY 13 CLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db 1 CLNGGKCVTKYFNSNQRCSCKPKFQGEHCEIDTSKTCYQNGHSTRKANRDLSGRPCL 60
QY 73 PWN SATVLOQTY-HAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMVHDC 132
Db 61 AWDSPVLLKMYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMVQDCS 120
QY 133 DGKFPSSPEELKFCQGQKTLRPRFKIIGBEFTTIEN 169
Db 121 VGKSPSSPREKBEFQCGQKALRPRFKIVGGQVTNAEN 157

RESULT 15
O35727
```

ID	Q35727	PRELIMINARY;	PRT;	597 AA.	
AC	Q35727;				
DT	01-JAN-1998	(TrEMBLrel. 05, Created)			
DT	01-JAN-1998	(TrEMBLrel. 05, Last sequence update)			
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)			
DE	Factor XII.				
GN	P12.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Liver;				
RA	Schloesser M., Schwager S., Engel W.;				
RL	Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.				
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.				
CC	-!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.				
DR	EMBL; X99571; CAA67891.1; -.				
DR	HSSP; P00760; 1AQ7.				
DR	MEROPS; S01.211; -.				
DR	MGI; MGI:1891012; F12.				
DR	InterPro; IPR001314; Chymotrypsin.				
DR	InterPro; IPR006209; EGF like.				
DR	InterPro; IPR000083; Fibnctnl.				
DR	InterPro; IPR000562; FN Type II.				
DR	InterPro; IPR006210; IEGF.				
DR	InterPro; IPR000001; Kringle.				
DR	InterPro; IPR001254; Ser_protease_Try.				
DR	Pfam; PF00008; EGF; 2.				
DR	Pfam; PF00039; fn1; 1.				
DR	Pfam; PF00040; fn2; 1.				
DR	Pfam; PF00051; kringle; 1.				
DR	Pfam; PF00089; trypsin; 1.				
DR	PRINTS; PR00722; CHYMOTRYPSIN.				
DR	PRINTS; PR00013; FNTYPEII.				
DR	PRINTS; PR00018; KRINGLE.				
DR	ProDom; PD000995; FN Type II; 1.				
DR	ProDom; PD000395; Kringle; 1.				
DR	SMART; SM00181; EGF; 2.				
DR	SMART; SM00058; FN1; 1.				
DR	SMART; SM00059; FN2; 1.				
DR	SMART; SM00130; KR; 1.				
DR	SMART; SM00020; Tryp_spc; 1.				
DR	PROSITE; PS00022; EGF_1; 2.				
DR	PROSITE; PS01186; EGF_2; 1.				
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.				
DR	PROSITE; PS00023; FIBRONECTIN_2; 1.				
DR	PROSITE; PS00021; KRINGLE_1; 1.				
DR	PROSITE; PS00070; KRINGLE_2; 1.				
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.				
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.				
DR	PROSITE; PS00135; TRYPSIN_SER; 1.				
KW	EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;				
KW	Serine protease.				
SQ	SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;				
	Query Match				
	Best Local Similarity 29.4%; Score 677; DB 11; Length 597;				
	Matches 149; Conservative 64; Mismatches 162; Indels 52; Gaps 10;				
QY	13 CLNGGTCVSNKYFSNIHWCNCPKFGGQCEIDKSKTCYEGNGHFGYRGKASTDTMGRPCL 72				
Db	183 CLNGGSL---LVEDHPLCRPTGYGYFCDDLWATCYEGRLSVRGAGTQSGAPCQ 239				
QY	73 PMSATVLOQTY-HAHRSDALGLGKHNYCRPNDRPCVQVGLKPLVQECMVHDC 131				
Db	240 RW----TVEATYRNMTKEQALSWGLGHAFCRPNDRPCVQVGLKPLVQECMVHDC 295				
QY	132 -----ADGKKPSPP-----EELKFCQCG---KTLRPRFKI 159				
Db	296 QPTTAPLVVPSQESQAPSLSHAPNDSDTHQTSLSKNTMTMGGCGQFRKGLSFMRV 355				
QY	160 IGGEFTTIENQPFALYRRHGRGSVTYVCGGSLISPCVWISATHCFIDYPKKEDYIVYL 219				
Db	356 VGLGVALPGSHPIYAAALWGN-----FCAGSLIAPCWLTAAHCLQNRPAPEELTVL 409				
QY	220 GRSRLNSNTQGENKFEVENLLHKDYSADTFLAHNDIALKIR-SKEGRCAQPSRTIQT 278				
Db	410 GQDRHNSCEWCQTLAVRSYRLHEGFSSITYQH--DLALLRLQESKTNSCAILSPHVQPV 467				
QY	279 CLFSMYNDPQFGTSCBITGFGKENSTDYLYPEOLKMTVVKLISHRECCQPHYGVSEVTK 338				
Db	468 CLFSGAAPPSETVLCVAGWGHLQEGAEFESTLQEAQVFFIALDRCSNSNVHGDALPG 527				
QY	339 MLCADPQWKTDSCQDSDGSLVLC---SLQGRWTLGTIVSWGRGCAKDKPGVYRVSHF 395				
Db	528 MLCAGFLEGGTDACQDSDGSLVLCBEGTABHQLTLRGVLSWGGCGDRNKPVGVTVDV 587				
QY	396 LPWIRSH 402				
Db	588 LAWIQKH 594				
	RESULT 16				
	Q8K0D2 PRELIMINARY; PRT; 517 AA.				
ID	Q8K0D2				
AC	Q8K0D2;				
DT	01-OCT-2002	(TrEMBLrel. 22, Created)			
DT	01-OCT-2002	(TrEMBLrel. 22, Last sequence update)			
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)			
DE	Hypothetical protein.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Kidney;				
RA	Strausberg R.;				
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.				
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.				
CC	-!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.				
DR	EMBL; BC031775; AAH31775.1; -.				
DR	InterPro; IPR001314; Chymotrypsin.				
DR	InterPro; IPR000742; EGF_2.				
DR	InterPro; IPR001881; EGF_Ca.				
DR	InterPro; IPR006209; EGF like.				
DR	InterPro; IPR006210; IEGF.				
DR	InterPro; IPR000001; Kringle.				
DR	InterPro; IPR001254; Ser_protease_Try.				
DR	Pfam; PF00008; EGF; 2.				
DR	Pfam; PF00051; kringle; 1.				
DR	Pfam; PF00089; trypsin; 1.				
DR	PRINTS; PR00722; CHYMOTRYPSIN.				
DR	PRINTS; PR00018; KRINGLE.				
DR	ProDom; PD000395; Kringle; 1.				
DR	SMART; SM00181; EGF; 2.				
DR	SMART; SM00058; FN1; 1.				
DR	SMART; SM00059; FN2; 1.				
DR	SMART; SM00130; KR; 1.				
DR	SMART; SM00020; Tryp_spc; 1.				
DR	PROSITE; PS00022; EGF_1; 2.				
DR	PROSITE; PS01186; EGF_2; 1.				
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.				
DR	PROSITE; PS00023; FIBRONECTIN_2; 1.				
DR	PROSITE; PS00021; KRINGLE_1; 1.				
DR	PROSITE; PS00070; KRINGLE_2; 1.				
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.				
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.				
DR	PROSITE; PS00135; TRYPSIN_SER; 1.				
KW	EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;				
KW	Serine protease.				
SQ	SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;				
	Query Match				
	Best Local Similarity 38.0%; Score 674.5; DB 11; Length 517;				
	Matches 159; Conservative 58; Mismatches 164; Indels 37; Gaps 16;				
QY	13 CLNGGTCVSNKYFSNIHWCNCPKFGGQCEIDKSKTCYEGNGHFGYRGKASTDTMGRPCL 72				

```
Db 116 CQGGVCSRRHRSRF--TCACPDQYKGFCEIGPD-DCYVGDGYSYRGKSKTVNQNPCL 173
QY 73 PWSATVLOQTYHAHRSADLQGLGKHNCRNPNRRRPWCYQVGLKPLVQE--CMVHDC 131
Db 174 YWNSHLLQETYNFMEDAEHTGIAEHFNCNPDGDHKPWCYKVNSEKVKWEYCDVTVC 233
QY 132 ADGKKPSPPEEL-----KFG-CQOKTLRPF--KIIGETTTIENQFWFAIY--- 177
Db 234 PVPDTP-NPVESLLEPVMELPFGFSCGKTEVAEHAVKRIYGGFKSTAGKHPQWVSLQTS 292
QY 178 ----RRHGGSVTVVCGSLISPCWISATHCFDYPKEDYIVYLGSRSLNSNTQCEMK 233
Db 293 PLTTSMPQG---HFCCGALHPCWVLTAAHC-TDINTKHLKV-LGDQDLKKTESHEQT 346
QY 234 FEVENLLHKDYSADTLAHNDIALLKIRSEKGRCAQPSRTIQTICLPSMYNDP-QFGTS 292
Db 347 FRVEKILKYSQYNERDEIPHNDIALLKLPVGGHCALESRYVKTVCLPS---DFPSTG 403
QY 293 CEITGFGKENSTLYPEOLKMTVVVKLISHRECQOPHYGSEVTTMLCAADPQWK-TDS 351
Db 404 CHISGWTGTGE--GSRQLLDAKVLIANPLCNSRLQYDHTIDDSMICAGNLQKPGSDT 461
QY 352 CQDSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHPVIRSHTRKEENG 409
Db 462 CQDSGGPLCEKDGTYVYVYVGVSWGQECG--KFGVYTVQTKFLNWKTMHREAGL 517

RESULT 17
QI4520 PRELIMINARY; PRT; 560 AA.
AC Q14520; O00663;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HGF activator like protein (Hyaluronan binding protein 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA Tomita M.;
RT "Purification and characterization of a novel hyaluronan-binding
protein (PHBP) from human plasma; it has three EGF, a kringle and a
serine protease domain, similar to hepatocyte growth factor
activator.",
RL J. Biochem. 119:1157-1165 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96425001; PubMed=8827452;
RA Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
RA Tomita M.;
RP SEQUENCE FROM N.A.
RA Kitamura N.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96425001; PubMed=8827452;
RA Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
RA Tomita M.;
RT "Purification and characterization of a novel hyaluronan-binding
protein (PHBP) from human plasma; it has three EGF, a kringle and a
serine protease domain, similar to hepatocyte growth factor
activator.",
RL J. Biochem. 119:1157-1165 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; D49742; BAA08576.1; -.
DR EMBL; S83182; AAB46909.1; -.
DR EMBL; BC031412; AAB31412.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.033; -.
DR Genew; HGNC:4798; HAP2.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00051; kringle; 1.
```

```
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;

Query Match 28.7%; Score 661.5; DB 4; Length 560;
Best Local Similarity 37.1%; Pred. No. 4.4e-55;
Matches 155; Conservative 58; Mismatches 166; Indels 39; Gaps 15;

QY 13 CLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRGKASTDTMTGRPCL 72
Db 159 CQNGATCSRHKRSKF--TCACPDQYKGFCEIG-SDDCYVGDGYSYRGKMNRTVNOHACL 216
QY 73 PWSATVLOQTYHAHRSADLQGLGKHNCRNPNRRRPWCYQVGLKPL-----VQEC 126
Db 217 YWNSHLLQENYANFMEDAEHTGIEHNFNCNPDADKPCWCFIKVTNDKVKWEYCDVSAC 276
QY 127 MVHDCA-DGKKPSPPEELK--FQCQOKTLRPF--FKIIGETTTIENQFWFAIYRRHR 181
Db 277 SAQDVAYPEESPTPESTKLPFGDSCGKTEAERKIRIYGGFKSTAGKHPQASLQ---- 332
QY 182 GGSVT-----YVCGSLISPCWISATHCFIDYPKEDYIVYLGSRSLNSNTQCEMK 233
Db 333 -SSLPLTISMPQGHFGGALIHPCWVLTAAHC-TDI-KTRHLKVVLGDQDLKKEEFHEQS 389
QY 234 FEVENLLHKDYSADTLAHNDIALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQF--GT 291
Db 390 FRVEKIFKYSHYNERDEIPHNDIALLKLPVGGHCALESRYVKTVCLP----DGSFPSSG 445
QY 292 SCEITGFGKENSTLYPEOLKMTVVVKLISHRECQOPHYGSEVTTMLCAADPQWK-TD 350
Db 446 ECHISGWTGTGK--GSRQLLDAKVLIANPLCNSRLQYDHTIDDSMICAGNLQKPGQD 503
QY 351 CQDSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHPVIRSHTRKEENG 408
Db 504 TCQDSGGPLTCEKDGTYVYVYVGVSWGLEG--KFGVYTVQTKFLNWKATIKSESG 559

RESULT 18
O97587 PRELIMINARY; PRT; 128 AA.
AC O97587;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand White;
RX MEDLINE=99057575; PubMed=9837780;
RA Reno C., Boykiw R., Martinez M.L., Hart D.A.;
RT "Temporal alterations in mRNA levels for proteinases and inhibitors
and their potential regulators in the healing medial collateral
ligament.",
RL Biochem. Biophys. Res. Commun. 252:757-763 (1998).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF069711; AAC95003.1; -.
DR HSSP; P00749; 1EJN.
DR MEROPS; S01.231; -.
```

```

DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Hydrolase; Kinase; Protease; Serine protease.
FT NON_TER 1 128
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 14328 MW; 1BC7ED30E071A06D CRC64;

Query Match 25.4%; Score 585; DB 6; Length 128;
Best Local Similarity 84.4%; Pred. No. 1.6e-48;
Matches 108; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Qy 172 WFAIYRHRGGSVTVYCGSLSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGE 231
Db 1 WFAIYRHRGGSVTVYCGSLSPCWVISATHCFINHQKEDYIVYLGSRSLNSNTPGE 60

Qy 232 MKFEVENLILHKDYSADTLAHNDIALKIRSEKRCQAQPSRTIQTICLPSMYNDPQFGT 291
Db 61 MKFEVEQLLHGYRADTLAHNDIALKILSNNGCAQPSRSITQTICLPPWNADPNFGT 120

Qy 292 SCEITGFG 299
Db 121 SCEITGFG 128

RESULT 19
Q91WJ5 PRELIMINARY; PRT; 812 AA.
AC Q91WJ5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen.
GN PLG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Brathwaite M., Waeltz P., Qian Y., Dudekula D., Schlessinger D.,
RA Nagaraja R.;
RT "Genomic Sequence Analysis in the Mouse t-complex Region.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; BC014773; AAH14773.1; -.
DR EMBL; AF481053; AAM22156.1; -.
DR HSSP; P00761; 1ANI.
DR MGD; MGL:97620; Plg.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003966; Prothrombin.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR001400; Somatostatin.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00118; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR Predom; PD000395; Kringle; 5.
DR PROSITE; PS00021; KRINGLE_1; 5.

DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00338; SOMATOSTATIN; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 812 AA; 90781 MW; 2417326086A2FFD2 CRC64;

Query Match 21.8%; Score 501; DB 11; Length 812;
Best Local Similarity 32.5%; Pred. No. 2.2e-39;
Matches 135; Conservative 48; Mismatches 170; Indels 62; Gaps 15;

Qy 10 NCDCLNGGTCVSNKYPSNIHWCNCPK--KFGQHQCEI-----DKSKTCVEGNHG 56
Db 428 NPDGDKGPWCYTTDPSVRWEYCNLRCSETGSGSVVLEPTVSGPSPGSDSETDCMYGNKG 487

Qy 57 FYRGKASDTMGRPCLPWNSATVLQOTVHAHRSALQ---LGLKHNCRNPD-NRRRP 111
Db 488 DIRGKTAVTAAGTTCQGWAA-----QEPHRSIFTPQTNPRAGLEK-NYCRNPDGDNVP 541

Qy 112 WCYVQVGLKPLVQECMVHDCADGKKPSPBPBELKQCQKTLRPR---FKIIGSBFTTIE 168
Db 542 WCYT-TNPKLYDYCDIPLCASAS-----SPECGKQPVEPKKCPGRVVGVCVANH 591

Qy 169 NOPWFAAIYRHRGGSVTVYCGSLSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNT 228
Db 592 SWPQWISLRTFTG---QHFCGGTLLAPEWLVTAACHLEKSSRPFEYKVLGAHEEYIRG 648

Qy 229 QGEMKFEVENLILHKDYSADTLAHNDIALKIRSEKRCQAQPSRTIQTICLPSMYNDPQ 288
Db 649 SDVQEISVAKLLE-----PNNRDIALKL-----SEPAITDKVIPACLFPS---PN 692

Qy 289 P-----GTSCEITGPKENSTLYPEQLKMTVVKLIHRECOQPHYSEVTTMLCAAD 344
Db 693 YMVADRTICYITGWTGTGT--FGAGRLKBAQLPIENKVCNRYEYLNVRVAKSTELCAGQ 750

Qy 345 PQWKTDSCQDGGGLVCSLQGRWTLTGIVSVGRGKALKDKPGVTVTRYSHFLPWI 399
Db 751 LAGVDSCQDGGGLVCSLQGRWTLTGIVSVGRGKALKDKPGVTVTRYSHFLPWI 805

RESULT 20
Q15146 PRELIMINARY; PRT; 810 AA.
AC Q15146;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RA Mitchell D., Robinson J.H.;
RT "Expression of recombinant human plasminogen and aglycoplasminogen in
RT HeLa cells.";
RL Fibrinolysis 0:0-0(1991).
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; M74220; AAA36451.1; -.
DR HSSP; P00747; 2PK4.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan-app.
DR InterPro: IPR003966; Prothrombin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.

```



```
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; Tryp SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 5.
DR PROSITE; PS50070; KRINGLE 2; 5.
DR PROSITE; PS02040; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN HIS; 1.
DR PROSITE; PS00135; TRYPSIN SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 810 PLASMINOGEN.
SQ SEQUENCE 810 AA; 90555 MW; B05C7D4B0D020B3C CRC64;

Query Match 21.6%; Score 498; DB 4; Length 810;
Best Local Similarity 33.3%; Pred. No. 4.2e-39;
Matches 138; Conservative 45; Mismatches 168; Indels 64; Gaps 16;

QY 10 NCDCNLGGTCVSNKYFNSIHNCPCPKFGGQ-----HCEIDSKTCYEGNGH 56
DB 428 NPDAKGPWCFTTDPVSRWBYCNLKKCGTEASVAPPVVLDPDVTPEEDCMFGNGK 487
QY 57 FYRGKASTDTMGRPCLPWNSATVLQOYYHAHR-----SDALQLGLGKHNYCRNPD-NRRRP 111
DB 488 GYRGKATTTGTGFCQDAA-----QEPHRHSIFTPETNPRAGLEK-NYCRNPDDGVGGP 541
QY 112 WCYVQGLKPLVQRCMVHDCADGKKPSPPEELKFCQCKTLRPR---FKIIGGEFTTIE 168
DB 542 WCYT-TNPRKLYDCVDFQCA---APS-----FDCGKQVPEKPCGPRVGGCVVAHPH 590
QY 169 NQWFAAIYRRHGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNT 228
DB 591 SNWQVSL--RTRFG--MHFCGGTLISPEWLTAAHCLKSPRPSSYYKVLGAHQEVNLE 646
QY 229 QGEMKFEVENLILHKDYSADTLAHHNDIALKIRSEKRCQAQPSRTTQICLPSMYNDPQ 288
DB 647 PHQVEIERSLFLPT-----RKDIALLKLSPP-----AVITDKVIPACLES----FN 690
QY 289 F-----GTSCEITGFKENSNDYLYPEOLKMTVVVKLISHRECQOHPHYGVSEVTTKMLCAAD 344
DB 691 YVVADETRCEFTWGNETQGT--FGAGLLKEAQLPVIENKVCNRYEFLNGRVQSTELCAGH 748
QY 345 POKNTDSCQDGGPLVCSLQGRMTLTGIVSWRGCAKDKKPGVYTRVSHFLPWI 399
DB 749 LAGTDSQCQDGGPLVCFKDKYILQGVTSWGLGCAKPNKPGVYVRSRFTWI 803

RESULT 21
O46506 PRELIMINARY; PRT; 454 AA.
AC O46506;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein a (Fragment).
GN BABAPOA.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RA Cox L.A., Jett C., Hixson J.B.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice Site Mutation is Associated with Deletion of a Single Exon in a Null Allele."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
```

```
DR EMBL; AF029691; AAB97886.1; -.
DR HSP; P00747; 2PK4.
DR MEROPS; S01.999; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 2.
DR PROSITE; PS00070; KRINGLE 2; 2.
DR PROSITE; PS02040; TRYPSIN DOM; 1.
DR PROSITE; PS00135; TRYPSIN SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
KW Serine protease.
FT NON_TER 1 1
FT SIGNAL 1 19 POTENTIAL.
SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;

Query Match 21.6%; Score 496; DB 6; Length 454;
Best Local Similarity 34.0%; Pred. No. 3.1e-39;
Matches 130; Conservative 52; Mismatches 132; Indels 68; Gaps 15;

QY 50 CYEGNGHYFGKASTDTMGRPCLPWNSATVLQ--QTVHAHRSDALQLGLGKHNYCRNPND 107
DB 102 CYHGDGQSVRGFSFTVTGRTCSQSWSMTPHQHKRTPENHPDGLTN-----NYCRNPDA 156
QY 108 RRPWCYVQVGLPLV--QECMVHDCAD-----GKKPSSPPEEL 144
DB 157 DTGFWCT---MDPSRWCEYCNLTRCSDTGTGVTPLTVIPIPSLEARSQASS----- 208
QY 145 KFCQGKTLRPR---FKIIGGEFTTIENQWFAAIYRRHGGSVTVYVCGGSLISPCWVIS 201
DB 209 -FDCGKQVPEKPCGPRVGGCVVAHSAHWPQVSL--RTRFGK--HFCGGLISPEWILT 263
QY 202 ATHCFIDYPKKEDIYVILGRSLNSNTQGMKEVENLILHKDYSADTLAHHNDIALKLI 261
DB 264 AARCLEMSRPSSYYKVLG-----AHQVNLSEHVOIEVSKLFSEPTGA---DIALKL 315
QY 262 RSKEGCAQPSRTTQICLPSMYNDPQF---GTSCEITGFKENSNDYLYPEOLKMTVV 317
DB 316 ----SPALITDKVIPACLES-----PNYVITAMTECIITGWGETGT--FGAGLLREARL 365
QY 318 KLISHRECQOHPHYGVSEVTTKMLCAADPQWKTDSQDGGPLVCSLQGRMTLTGIVSWG 377
DB 366 PVIENTVCNRYEFLNGRVKSTELCAGHLAGTDSQCQDGGPLVCFKDKYILRGITSWG 425
QY 378 RGCALKDKKPGVYTRVSHFLPWI 399
DB 426 PGCAEPNKGIVYVRSRFTWI 447

RESULT 22
Q9ROW3 PRELIMINARY; PRT; 812 AA.
ID Q9ROW3
AC Q9ROW3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bangert K., Johnsen A.H., Thorsen S.;
RT "Rat plasminogen: cDNA and gene structure.";
```

Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RESULT 23

Q95EF3 PRELIMINARY; PRT; 300 AA.

AC Q95EF3

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Hypothetical protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

SEQUENCE FROM N.A.

RC TISSUE=Lung;

RA Strausberg R.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

DR EMBL; BC012390; AAH12390.1; -.

DR HSSP; P00761; IANI.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR001254; Ser. protease\_Try.

DR PRINTS; PR00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; Tryp\_SPC; 1.

DR PROSITE; PS0240; TRYPsin\_DOM; 1.

DR PROSITE; PS00134; TRYPsin\_HIS; 1.

DR PROSITE; PS00135; TRYPsin\_SER; 1.

KW Hypothetical protein; Hydrolase; Protease; Serine protease.

SQ SEQUENCE 300 AA; 32216 MW; FFC2BDF9382F636A CRC64;

Query Match 21.5%; Score 495.5; DB 4; Length 300;

Best Local Similarity 38.4%; Pred. No. 2.1e-39; Mismatches 112; Indels 15; Gaps 5;

Matches 107; Conservative 45;

Qy 132 ADGKSPPEELKFCQGG---KTLRPFKIKIGGEFTTIENQFWFAAIYRRHGGSVTV 188

Db 28 AKRQPPSLTRNGPLSCGQRLKSLSSMTRVVGLVALGAHPYIAALYMGHS-----F 81

Qy 189 CGSLISPCWISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSAD 248

Db 82 CAGSLIAPCWLTAAHCLQDRPAEDLTIVLQGRNHSCEPCQTLAVRSYRLHEAFS-- 139

Qy 249 TLAHNDIALLKIR-SKEGRCAQPSRTIOTICLPSMYNDPQFGTSCETITGCKENSTDYL 307

Db 140 PVSQHDALLRLQEDADGSCALLSYQVCLPFGAARPSSETTLQVAGMHQEGAE 199

Qy 308 YPEQLKMTVVKLISHRECQPHYGYSEVTTKMLCAADPQWKTDSCQSGGSLVCSLQ- 366

Db 200 YASFLQEAQVPELSLRCSPADPVHGSILPGLMCAAGLEGGTDACQDGGSLVCEQAA 259

Qy 367 --RMTLTGVISWGRGKALKKQPVYTRVSHFLPWIRSH 403

Db 260 ERLTLQGIISWGGCGDRNKPQVITDVAIYLAWIREHT 298

RESULT 24

Q95M89 PRELIMINARY; PRT; 103 AA.

AC Q95M89

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Plasminogen activator urokinase (Fragment).

GN PLAU.

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OX NCBI\_TaxID=9796;

RN [1]

SEQUENCE FROM N.A.

RC MEDLINE=21314992; PubMed=11421942;

RX Shubitowski D.M., Ventr P.J., Douglass C.L., Zhou R.-X., Ewart S.L.;

Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

RESULT 23

Q95EF3 PRELIMINARY; PRT; 300 AA.

AC Q95EF3

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Hypothetical protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

SEQUENCE FROM N.A.

RC TISSUE=Lung;

RA Strausberg R.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

DR EMBL; BC012390; AAH12390.1; -.

DR HSSP; P00761; IANI.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR001254; Ser. protease\_Try.

DR PRINTS; PR00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR SMART; SM00020; Tryp\_SPC; 1.

DR PROSITE; PS0240; TRYPsin\_DOM; 1.

DR PROSITE; PS00134; TRYPsin\_HIS; 1.

DR PROSITE; PS00135; TRYPsin\_SER; 1.

KW Hypothetical protein; Hydrolase; Protease; Serine protease.

SQ SEQUENCE 300 AA; 32216 MW; FFC2BDF9382F636A CRC64;

Query Match 21.5%; Score 495.5; DB 4; Length 300;

Best Local Similarity 38.4%; Pred. No. 2.1e-39; Mismatches 112; Indels 15; Gaps 5;

Matches 107; Conservative 45;

Qy 132 ADGKSPPEELKFCQGG---KTLRPFKIKIGGEFTTIENQFWFAAIYRRHGGSVTV 188

Db 28 AKRQPPSLTRNGPLSCGQRLKSLSSMTRVVGLVALGAHPYIAALYMGHS-----F 81

Qy 189 CGSLISPCWISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSAD 248

Db 82 CAGSLIAPCWLTAAHCLQDRPAEDLTIVLQGRNHSCEPCQTLAVRSYRLHEAFS-- 139

Qy 249 TLAHNDIALLKIR-SKEGRCAQPSRTIOTICLPSMYNDPQFGTSCETITGCKENSTDYL 307

Db 140 PVSQHDALLRLQEDADGSCALLSYQVCLPFGAARPSSETTLQVAGMHQEGAE 199

Qy 308 YPEQLKMTVVKLISHRECQPHYGYSEVTTKMLCAADPQWKTDSCQSGGSLVCSLQ- 366

Db 200 YASFLQEAQVPELSLRCSPADPVHGSILPGLMCAAGLEGGTDACQDGGSLVCEQAA 259

Qy 367 --RMTLTGVISWGRGKALKKQPVYTRVSHFLPWIRSH 403

Db 260 ERLTLQGIISWGGCGDRNKPQVITDVAIYLAWIREHT 298

RESULT 24

Q95M89 PRELIMINARY; PRT; 103 AA.

AC Q95M89

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Plasminogen activator urokinase (Fragment).

GN PLAU.

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OX NCBI\_TaxID=9796;

RN [1]

SEQUENCE FROM N.A.

RC MEDLINE=21314992; PubMed=11421942;

RX Shubitowski D.M., Ventr P.J., Douglass C.L., Zhou R.-X., Ewart S.L.;

RT "polymorphism identification within 50 equine gene-specific sequence  
tagged sites";

RL Anim. Genet. 32:78-78(2001).  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
DR EMBL: AY008806; AAK14840.1; .  
DR EMBL: AY008803; AAK14840.1; JOINED.  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR001254; Ser\_protease\_Try.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
KW Hydrolase; Kinase; Protease; Serine protease.  
FT NON\_TER 1  
FT NON\_TER 103  
SQ SEQUENCE 103 AA; 11525 MW; 0B739514F6331180 CRC64;

Query Match 21.5%; Score 495; DB 6; Length 103;  
Best Local Similarity 87.4%; Pred. No. 5.9e-40;  
Matches 90; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
QY 172 WFAAIYRRHRGGSVTVCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGE 231  
Db 1 WFAAIYRRHRGGSVTVCGSLISPCWVLSATHCFINYPKEDYIVYLGSRSLNSTSGE 60  
QY 232 MKFEVENLILHKDYSADTLAHNDIALLKIRSGRCQAQPSRT 274  
Db 61 MKFEVKLIILHEDYSADTLAHNDIALLKITSSTGTGCAQPSRS 103

## RESULT 25

Q8AVB0 PRELIMINARY; PRT; 429 AA.  
AC Q8AVB0  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DE Plasminogen precursor (Fragment).  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hanumanthaiah R., Day K., Jagadeeswaran P.;  
RT "Comprehensive analysis of blood coagulation pathways in teleostei:  
RT Evolution of coagulation factor genes and identification of zebrafish  
RT Factor VIIi.";  
RL Blood Cells Mol. Dis. 0:0-0(2002).  
DR EMBL; AF515276; AAN71006.1; .  
FT NON\_TER 1  
SQ SEQUENCE 429 AA; 47556 MW; 9A580A214A549C12 CRC64;

Query Match 21.4%; Score 491.5; DB 13; Length 429;  
Best Local Similarity 35.3%; Pred. No. 8e-39;  
Matches 130; Conservative 36; Mismatches 153; Indels 49; Gaps 14;  
QY 48 KTCYEGNGHFYRGKASDTDMGRCLPWNSTAVLQQ---TVHAUSDALQLGLGHKHYCRN 104  
Db 100 KCKNGNGAFYRGSTSMTVGVTCQAWRSMTPHQASFTPTHDPKRGLE-----SNQCRN 154  
QY 105 PDNR-RRPWCYVQVGLKPLVQECWHDCAADGKPKSPPEELKFCQCKTLRPR---FKII 160  
Db 155 PDSVNGPWCYTTDPSKKW-DYCIQIDP-----ESLK--CQPKATPKRCRGRIV 201  
QY 161 GGEFTTIENQWPFAAIYRRHRGGSVTVCGSLISPCWVISATHCFIDYPKEDYIVYLG 220  
Db 202 GGCVSXPHSPWP--QISLRTRG--KIHFCGGTLIDPQWVTAHCLERSDPSAYKIMLG 257  
QY 221 RSLNSNTQGEKMFVENLILHKDYSADTLAHNDIALLKIRSGRCQAQPSRTIOTICL 280

Db 258 IHTERATESKQBRDVTKII---KGPAQT-----DIALKL-----DRPALINDKVSPLVCL 305  
QY 281 PSMYNDPQFGTSCETITGFGKENST---DYLYPQLKMTVVKLISHRECOQPHYGVSEVTT 337  
Db 306 PEKDYIVPSNTECVVTGWGETQDTGGEGY-----LKETGPPVIENKVCNRPFLNGRVKD 360  
QY 338 KMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSGRGCALDKDKPGVYTVRSHPLP 397  
Db 361 HENCAGNIEGGNDSCQDGGPLVCAQNTFVLQGVTSWGLGCANAMKPGVYTVRSKFEVD 420  
QY 398 WIRSHTKE 405  
Db 421 WIRSKE 428

## RESULT 26

O46507 PRELIMINARY; PRT; 334 AA.  
AC O46507  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Plasminogen (Fragment).  
GN BABPSPSG  
OS Papio hamadryas (Hamadryas baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Butheria; Primates; Catarrhini; Cercopitheidae;  
OC Cercopitheidae; Papio.  
OX NCBI\_TaxID=9557;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Cox L.A., Jett C., Hixson J.E.;  
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice  
RT Site Mutation is Associated with Deletion of a Single Exon in a Null  
RT Allele."  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; AF029692; AAB97887.1; .  
DR HSSP; P00747; SHPG.  
DR MEROPS; S01.233; .  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003966; Prothrombin.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR01505; PROTHROMBIN.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.  
FT NON\_TER 1  
SQ SEQUENCE 334 AA; 36791 MW; C7DC06B03B965286 CRC64;

Query Match 21.0%; Score 482.5; DB 6; Length 334;  
Best Local Similarity 35.4%; Pred. No. 4.3e-38;  
Matches 129; Conservative 46; Mismatches 134; Indels 55; Gaps 16;  
QY 50 CYEGNGHFYRGKASDTDMGRCLPWNSTAVLQOTVHAH---RSDALQLGLGHKHYCRNP 105  
Db 5 CMFGNGKRYRGKATVTVGTPOEWA---KEPHSLIFTPTTPRAGLEK-NYCRNP 59  
QY 106 D-NRRRPWCYVQVGLKPLVQECWHDCAADGKPKSPPEELKFCQCKTLRPR---FKIIG 161  
Db 59 DGDVGSGPWCYT-TNPRKLYDYCDVPCASS-----SFDCKGPKQVEPKKCPGRVVG 107

```

QY 162 GEFTIENQWPAALYRRHGGSVYVCGGSLISPCWVISATHCFIDYPKKEDYIVVLGR 221
Db 108 GCVAHAHSWQVSL--RTEFG--MHECGGTLISPEWVLTAAHLEKSPSPSYKILGA 163
QY 222 SRLNSNTQGMKFF--VENILHKDYSGADTLAHNDIALKIRSKRGCACQAPORTIQTIC 279
Db 164 HQ-----EVRLPHVQIEVSKMFSEPAGA---DIALKLSP-----AIIIDKVIAPAC 209
QY 280 LPSMYNDPQF-----GTCEITGFKENSTLYPEQLKMTVVKVLISHRECCQPHYVSEV 335
Db 210 LPS-----PNYVADRTECFITGGETGT--YGAGLKEARLPVIEKNVCRVEFLNGRV 263
QY 336 TTKMLCAADPOWKTDCQGGSGPLVCSLQGRMTLTGIVSWRGCCALKDKPGYVTRVSHF 395
Db 264 KSTELCAGHLAGGDSQGGSGPLVCFEKDKYILQGVTSWGLGCARPKNKPGYVVRVSRF 323
QY 396 LPWI 399
Db 324 VTWI 327

RESULT 27
Q18783
ID O18783 PRELIMINARY; PRT; 806 AA.
AC O18783;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen.
OS Macropus eugenii (Tamar wallaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98004511; PubMed=9342350;
RA Lawn R.M., Schwartz K., Parthy L.;
RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).
CC -/- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AF012297; AAB65760.1; -.
DR HSP; P00747; SHPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TRY_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS0240; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin HIS; 1.
DR PROSITE; PS00135; TRYPsin SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 806 AA; 90981 MW; 95FAA86DC20064D5 CRC64;

```

Query Match 20.5%; Score 472.5; DB 6; Length 806;  
 Best Local Similarity 32.0%; Pred. No. 1.2e-36;  
 Matches 132; Conservative 43; Mismatches 162; Indels 75; Gaps 16;

```

QY 13 CLNGTCVSNKYFSNIHWCNCPKXFGGQHCBIDSKTCYENGHGYRCKASTDDTMGRPCL 72
Db 448 CSGTGSTVLNAQTTRV-----PSVDTTSHPSD-----CMYSGKDYRGKRTTGTGLCQ 498
QY 73 PWSATVLQOTVHAH-----RSDALQLGLCKHVCNRPD--NRRRPWCYVQVGLKPLVQECMV 128
Db 499 AWTA-----QEPHRTITPTDTPYPRAGLEENVCNRPDGPNGPWCYT--TNPKKLEFDYCDI 552
QY 129 HPCADGKPKSSPEBELKFCQCKTLRPR---FKIIGGEFTTIENQWPAALYRRHGGSV 185
Db 553 PQCV-----FDCGKPRVEPKCPGRIVGGCYAOPHSWPM--QISLRTFGE- 600
QY 186 TVVCGSLISPCWV:SATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENILHKDY 245
Db 601 -HFCGGTLIAPOWLVLTAAHCLERSQWPGAYKVLG-----LHREV 639
QY 246 SADTLAHH-----NDIALKIRSKRGCACQAPSRITQTICLPSMYNDPQFGTSC 293
Db 640 NPESYSQIEGVSRFLPKGLAADIALLKL-----NRPAINDKVIPACLFSDQPMVDRTLIC 695
QY 294 RITGFKENSTLYPE-OLKMTVVKVLISHRECCQPHYVSEVTTKMLCAADPOWKTDCS 352
Db 696 HVTGWGDTQGTG---PRGLLKQASLPEVIDNRVCNREHYINGRVKSTELCAGHLVGRGDS 752
QY 353 QGDSGGLVCSLQGRMTLTGIVSWRGCCALKDKPGYVTRVSHFLPWIRSHTK 404
Db 753 QGDSGGLICFEDDKVYLGQVTSWGLGCARPKNKPGYVVRVSRYSIWIWEDVMK 804

RESULT 28
Q9Y1V3
ID Q9Y1V3 PRELIMINARY; PRT; 868 AA.
AC Q9Y1V3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tunicate retinoic acid-inducible modular protease precursor.
GN TRAMP.
OS Polyandrocarpa misakiensis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Styelidae; Polyandrocarpa.
OX NCBI_TaxID=7723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=white spot;
RX MEDLINE=99423646; PubMed=10491255;
RA Ohashi M., Kawamura K., Fujii N., Yubisui T., Fujiwara S.;
RT "A retinoic acid-inducible modular protease in budding ascidians.";
RL Dev. Biol. 214:38-45(1999).
CC -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB030007; BAA82522.1; -.
DR HSP; P00763; IDPO.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00057; ldl_recept_a; 3.
DR Pfam; PF00530; SRCR; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00192; LDLA; 3.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00202; SR_2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01209; LDLRA_1; 3.
DR PROSITE; PS00068; LDLRA_2; 3.
DR PROSITE; PS0287; SRCR_2; 2.
DR PROSITE; PS0240; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.

```

```
KW Hydrolase; Protease; Serine protease; Signal.
FT SIGNAL 1 28 POTENTIAL.
SQ SEQUENCE 868 AA; 97660 MW; F71462865F36A6CA CRC64;

Query Match      20.1%; Score 463.5; DB 5; Length 868;
Best Local Similarity 38.9%; Pred. No. 9.9e-36;
Matches 107; Conservative 44; Mismatches 103; Indels 21; Gaps 10;

QY 137 PSSPPELKKQCCKTL-----RPFKLIIGEFFTIENQWPFAAIYRRHGGSVTVVCGGS 192
   |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 601 PTPPPMP--ECGRKFVIEAPLEPTARIVGSGSGTEPHENPWQAGIWL-----PWTYWC 652

QY 193 LISPCWVISATHCFI-DYPKKEDYIVVLGRSLNNTQGEMKFEVENILHKDYSADTILA 251
   |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 653 LIHPCWLTAAHCVFREYPIR-DYTIRLDGHITGVDDETEQLFKAEIIKH-DYNVT-- 708

QY 252 HHNDIALLKTRSKRGCAQPSRTIOTICLPMSYNDPFGTSCEITGFGENSTDY-LYPE 310
   |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 709 KENDIALLRINDARECATITPVQVTCVLFKSSQFDAKTICEVTGWGKDSATAVRAYVP 768

QY 311 QLKMTVVKLISHRECQQPHYVGSEVTTMLCAADPOWKTDSCQSGGGLVCSLG--RM 368
   |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 769 VLQAEIPLANKKCLRDSEY-TQLGPTFCAGYLTCGKSDSCQSGGGLPLCRDSDRY 827

QY 369 TLTGISVWGRCALKDKPGVYTRVSHFLPWIRSH 403
   |||||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 828 YVMGIVSWGNGCAKPAKVGYAVFIDWIEQM 862

RESULT 29
Q8NF86 PRELIMINARY; PRT; 327 AA.
AC Q8NF86;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Similar to protease, serine, 8 (Prostasin) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Darrow A.L., Qi J., Andrade-Gordon P., Chen C.;
RT "DNA encoding the human serine protease EOS.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF536382; AAN04055.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYP SIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 327 AA; 35106 MW; 60458EDB817AC1CF3 CRC64;

Query Match      19.2%; Score 442.5; DB 4; Length 327;
Best Local Similarity 38.8%; Pred. No. 3e-34;
Matches 104; Conservative 35; Mismatches 96; Indels 33; Gaps 8;

QY 148 CGOKTLRPFKLIIGEFFTIENQWPFAAIYRRHGGSVTVVCGSLISPCWVISATHCFI 207
   |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 75 CGQPRMSSR--IVGGRDGRDGWPWQASI--QHKG---AHVCGSLIAPQWLVTAACHFP 127

QY 208 DYPKKEDIYVLGRSLNNTQGEMKFEVENILHKDYSADTLAHHNDIALLKIRSKRGR 267
   |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 128 RRALPAEYRVRLGALRGSTSPRTLSPVRRVLLPDYSED--GARGDLALLQLR----R 181

QY 268 CAQPSRTIOTICLPMSYNDPFGTSCEITGFGENSTDYLP-----EQIKMTVKLI 320
   |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:43:04 ; Search time 42.1951 Seconds  
(without alignments)  
1811.566 Million cell updates/sec

Title: US-09-880-503-3

Perfect score: 2301

Sequence: 1 SNEHQVPSNCDCLNGTGV.....VSHFLPWIRSHYKXENGLAL 411

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2301	100.0	411	10	US-09-880-503-3
2	2301	100.0	431	12	US-10-301-822-161
3	2301	100.0	431	12	US-10-131-985-21
4	2301	100.0	431	14	US-10-076-421-2
5	2301	100.0	431	15	US-10-171-311-184
6	2299	99.9	431	12	US-10-247-671-149
7	2298	99.9	431	15	US-10-193-656-4
8	2291	99.6	411	12	US-10-407-821-2
9	2283	99.2	431	10	US-09-264-468B-1
10	2243	97.5	403	10	US-09-880-503-6
11	1737	75.5	323	10	US-09-880-503-7
12	1535	66.7	337	15	US-10-106-698-6266
13	1508	65.5	276	10	US-09-880-503-5
14	1464	63.6	268	12	US-10-407-821-3
15	1333	57.9	246	10	US-09-264-468B-2

16	1318	57.3	241	11	US-09-898-837A-47	Sequence 47, Appl
17	864.5	37.6	527	11	US-09-987-457-18	Sequence 18, Appl
18	864.5	37.6	527	11	US-09-987-455-19	Sequence 19, Appl
19	864.5	37.6	562	9	US-09-969-271-7	Sequence 7, Appl
20	864.5	37.6	562	10	US-09-974-298-145	Sequence 145, App
21	864.5	37.6	562	12	US-10-443-701-4	Sequence 4, Appl
22	864.5	37.6	562	15	US-10-193-656-8	Sequence 8, Appl
23	837	36.4	143	10	US-09-880-503-8	Sequence 8, Appl
24	793	34.5	135	10	US-09-880-503-4	Sequence 4, Appl
25	793	34.5	138	15	US-09-984-186-12	Sequence 12, Appl
26	793	34.5	138	15	US-10-237-667-12	Sequence 12, Appl
27	793	34.5	138	15	US-10-237-708-12	Sequence 12, Appl
28	793	34.5	138	15	US-10-237-866-12	Sequence 12, Appl
29	793	34.5	138	15	US-10-237-871-12	Sequence 12, Appl
30	793	34.5	138	15	US-10-237-624-12	Sequence 12, Appl
31	785.5	34.1	372	9	US-09-084-491A-3	Sequence 3, Appl
32	785.5	34.1	372	14	US-10-102-704-3	Sequence 3, Appl
33	780	33.9	354	11	US-09-987-457-10	Sequence 10, Appl
34	780	33.9	354	11	US-09-987-455-11	Sequence 11, Appl
35	780	33.9	377	11	US-09-987-455-8	Sequence 8, Appl
36	741	32.2	343	11	US-09-987-457-14	Sequence 14, Appl
37	741	32.2	343	11	US-09-987-455-15	Sequence 15, Appl
38	740	32.2	339	11	US-09-987-457-12	Sequence 12, Appl
39	740	32.2	339	11	US-09-987-455-13	Sequence 13, Appl
40	737	32.0	331	11	US-09-987-457-11	Sequence 11, Appl
41	737	32.0	331	11	US-09-987-455-12	Sequence 12, Appl
42	737	32.0	343	11	US-09-987-457-15	Sequence 15, Appl
43	737	32.0	343	11	US-09-987-455-16	Sequence 16, Appl
44	735.5	32.0	655	15	US-10-172-712-28	Sequence 28, Appl
45	733	31.9	335	11	US-09-987-457-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-880-503-3  
; Sequence 3, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; TITLE OF INVENTION: TISSUE CONTRACTILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880, 503  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-3

Query Match	100.0%	Score 2301;	DB 10;	Length 411;
Best Local Similarity	100.0%	Pred. No. 2.5e-194;		
Matches 411;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG	60	
DB	1	SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG	60	
QY	61	KASTDTMGPCLPWNSATVLOQTYHAHRSALQGLGKKNYCRNPNRRPWCYVOVGLK	120	
DB	61	KASTDTMGPCLPWNSATVLOQTYHAHRSALQGLGKKNYCRNPNRRPWCYVOVGLK	120	
QY	121	PLVQECWHDCAADGKKPSPPEELKFCQCKTLRPRFKLIIGGFTTIENQPFAYRRH	180	
DB	121	PLVQECWHDCAADGKKPSPPEELKFCQCKTLRPRFKLIIGGFTTIENQPFAYRRH	180	

```
QY 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 240
Db 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 240
QY 241 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 300
Db 241 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 300
QY 301 ENSTDYLYPEQLKMTVVKLIHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDSSGGL 360
Db 301 ENSTDYLYPEQLKMTVVKLIHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDSSGGL 360
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
Db 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 2
US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-029PRNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match 100.0%; Score 2301; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.7e-194;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGTGVSNKYFNSIHWNCNPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 21 SNEHQVPSNCDCLNGTGVSNKYFNSIHWNCNPKFGGQHCIDKSKTCYEGNGHYRG 80
QY 61 KASTDTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNYCRNPONRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNYCRNPONRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPKSPPEELKFCQCGOKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
Db 141 PLVQECMVHDCADGKPKSPPEELKFCQCGOKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200
QY 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 240
Db 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 260
QY 241 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 300
Db 261 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLIHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDSSGGL 360
Db 321 ENSTDYLYPEQLKMTVVKLIHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 4
US-10-076-421-2
; Sequence 2, Application US/10076421
```

```
QY 301 ENSTDYLYPEQLKMTVVKLIHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDSSGGL 360
Db 321 ENSTDYLYPEQLKMTVVKLIHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 3
US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Ocleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-21

Query Match 100.0%; Score 2301; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.7e-194;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGTGVSNKYFNSIHWNCNPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 21 SNEHQVPSNCDCLNGTGVSNKYFNSIHWNCNPKFGGQHCIDKSKTCYEGNGHYRG 80
QY 61 KASTDTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNYCRNPONRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNYCRNPONRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPKSPPEELKFCQCGOKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
Db 141 PLVQECMVHDCADGKPKSPPEELKFCQCGOKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200
QY 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 240
Db 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 260
QY 241 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 300
Db 261 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLIHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDSSGGL 360
Db 321 ENSTDYLYPEQLKMTVVKLIHRECCQPHYVGGSEVTTKMLCAADPQWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 4
US-10-076-421-2
; Sequence 2, Application US/10076421
```

```
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, NAOKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAVAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-2

Query Match      100.0%; Score 2301; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.7e-194;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 81 KASDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPSPEELKFCQGGKTLRPFKIIGGEFTTIENQPFPAIYRRH 180
DB 141 PLVQECMVHDCADGKPSPEELKFCQGGKTLRPFKIIGGEFTTIENQPFPAIYRRH 200
QY 181 RGSSTVTVCGSGLSPCWVISATHCFIDYPKKEDYIYVLSRLNSNTQGMKFEVENLI 240
DB 201 RGSSTVTVCGSGLSPCWVISATHCFIDYPKKEDYIYVLSRLNSNTQGMKFEVENLI 260
QY 241 LHKDYSADTLAHNDIALKIRSGRCAQPSRTIQICLPSPMYNDPFGTSCBITGFGK 300
DB 261 LHKDYSADTLAHNDIALKIRSGRCAQPSRTIQICLPSPMYNDPFGTSCBITGFGK 320
QY 301 ENSTDVILYPLQKMTVVKLISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 360
DB 321 ENSTDVILYPLQKMTVVKLISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALDKDQGVYTRVSHFLPWIRSHTKENGLAL 411
DB 381 VCSLQGRMTLTGIVSWGRCALDKDQGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 5
US-10-171-311-184
; Sequence 184, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersch, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159

; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, NAOKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAVAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-184

Query Match      100.0%; Score 2301; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.7e-194;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 81 KASDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPSPEELKFCQGGKTLRPFKIIGGEFTTIENQPFPAIYRRH 180
DB 141 PLVQECMVHDCADGKPSPEELKFCQGGKTLRPFKIIGGEFTTIENQPFPAIYRRH 200
QY 181 RGSSTVTVCGSGLSPCWVISATHCFIDYPKKEDYIYVLSRLNSNTQGMKFEVENLI 240
DB 201 RGSSTVTVCGSGLSPCWVISATHCFIDYPKKEDYIYVLSRLNSNTQGMKFEVENLI 260
QY 241 LHKDYSADTLAHNDIALKIRSGRCAQPSRTIQICLPSPMYNDPFGTSCBITGFGK 300
DB 261 LHKDYSADTLAHNDIALKIRSGRCAQPSRTIQICLPSPMYNDPFGTSCBITGFGK 320
QY 301 ENSTDVILYPLQKMTVVKLISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 360
DB 321 ENSTDVILYPLQKMTVVKLISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALDKDQGVYTRVSHFLPWIRSHTKENGLAL 411
DB 381 VCSLQGRMTLTGIVSWGRCALDKDQGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 6
US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match      99.9%; Score 2299; DB 12; Length 431;
```





QY 361 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTEENGLAL 411  
Db 361 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTEENGLAL 411

RESULT 9  
US-09-264-468B-1  
; Sequence 1, Application US/09264468B  
; Patent No. US20020106775A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Jieyi  
; APPLICANT: Nienaber, Vicki L.  
; APPLICANT: Henkin, Jack  
; APPLICANT: Smith, Richard A.  
; APPLICANT: Walter, Karl A.  
; APPLICANT: Severin, Jean M.  
; APPLICANT: Edalji, Rohinton  
; APPLICANT: Johnson Jr., Robert W.  
; APPLICANT: Holzman, Thomas F.  
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE  
; FILE REFERENCE: 6310.US.PL  
; CURRENT APPLICATION NUMBER: US/09/264,468B  
; PRIOR FILING DATE: 1999-03-05  
; PRIOR APPLICATION NUMBER: US 09/036,361  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(20)  
; OTHER INFORMATION: Leader sequence  
; NAME/KEY: VARIANT  
; LOCATION: (279)...(279)  
; OTHER INFORMATION: Xaa = any amino acid  
; NAME/KEY: VARIANT  
; LOCATION: (302)...(302)  
; OTHER INFORMATION: Xaa = any amino acid  
US-09-264-468B-1

Query Match 99.2%; Score 2283; DB 10; Length 431;  
Best Local Similarity 99.5%; Pred. No. 1e-192;  
Matches 409; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 SNELHQPNSCDCLNGGTCVSNKYFNSIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60  
Db 21 SNELHQPNSCDCLNGGTCVSNKYFNSIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 80  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120  
Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDRRRPWCYVQVGLK 140  
QY 121 PLVQECMVHDCADGKPPSPPELKFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 180  
Db 141 PLVQECMVHDCADGKPPSPPELKFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200  
QY 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 240  
Db 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 260  
QY 241 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCITGFGK 300  
Db 261 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCITGFGK 320  
QY 301 ENSTDYLYPEQLKMTVVKLIISHRECCQPHYVGSVTTKMLCAADPQWKTDSQGDGGPL 360  
Db 321 EXSTDYLYPEQLKMTVVKLIISHRECCQPHYVGSVTTKMLCAADPQWKTDSQGDGGPL 380  
QY 361 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTEENGLAL 411

Db 381 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTEENGLAL 431

RESULT 10  
US-09-880-503-6  
; Sequence 6, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 403  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-6  
Query Match 97.5%; Score 2243; DB 10; Length 403;  
Best Local Similarity 98.1%; Pred. No. 3.2e-189;  
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;  
QY 1 SNELHQPNSCDCLNGGTCVSNKYFNSIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60  
Db 1 SNELHQPNSCDCLNGGTCVSNKYFNSIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120  
Db 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGKPPSPPELKFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 180  
Db 121 PLVQECMVHDCADGKPPSPPELKFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 172  
QY 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 240  
Db 173 RGGSVTVVCGGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLI 232  
QY 241 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCITGFGK 300  
Db 233 LHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCITGFGK 292  
QY 301 ENSTDYLYPEQLKMTVVKLIISHRECCQPHYVGSVTTKMLCAADPQWKTDSQGDGGPL 360  
Db 293 ENSTDYLYPEQLKMTVVKLIISHRECCQPHYVGSVTTKMLCAADPQWKTDSQGDGGPL 352  
QY 361 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTEENGLAL 411  
Db 353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTEENGLAL 403

RESULT 11  
US-09-880-503-7  
; Sequence 7, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; PRIOR FILING DATE: 2000-06-20

```
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-7

Query Match      75.5%; Score 1737; DB 10; Length 323;
Best Local Similarity 78.6%; Pred. No. 8.8e-145;
Matches 323; Conservative 0; Mismatches 0; Indels 88; Gaps 1;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHYRG 60
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHYRG 47
Qy 61 KASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVLK 120
Db 48 -----
Qy 121 PLVQECMVHDCADGKFPSSPEELKFCQGQKTLRPRFKIIGGFTTIIENQPFWFAAIYRRH 180
Db 48 -----
Qy 181 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDIIVYLGRSRLNSNTQGMKEFEVENLI 240
Db 93 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDIIVYLGRSRLNSNTQGMKEFEVENLI 152
Qy 241 LHKDYSADTLAHNDIALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 300
Db 153 LHKDYSADTLAHNDIALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 212
Qy 301 ENSTDYLYPQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPL 360
Db 213 ENSTDYLYPQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPL 272
Qy 361 VCSLQGRMTLTIGVSGRGCKALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
Db 273 VCSLQGRMTLTIGVSGRGCKALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 323

RESULT 12
US-10-106-698-6266
; Sequence 6266, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6266
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-6266

Query Match      66.7%; Score 1535; DB 15; Length 337;
Best Local Similarity 95.2%; Pred. No. 5.7e-127;
Matches 275; Conservative 3; Mismatches 5; Indels 6; Gaps 1;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHYRG 60
Db 27 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHYRG 86
```

```
Qy 61 KASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVLK 120
Db 87 KASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVLK 146
Qy 121 PLVQECMVHDCADGKFPSSPEELKFCQGQKTLRPRFKIIGGFTTIIENQPFWFAAIYRRH 180
Db 147 PLVQECMVHDCADGKFPSSPEELKFCQGQKTLRPRFKIIGGFTTIIENQPFWFAAIYRRH 206
Qy 181 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDIIVYLGRSRLNSNTQGMKEFEVENLI 240
Db 207 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDIIVYLGRSRLNSNTQGMKEFEVENLI 266
Qy 241 LHKDYSADTLAHNDIALLKIRSEGRCAQ-----PSRTIQTICLPSM 283
Db 267 LHKDYSADTLAHNDIALLKIRSEGRCAQHPGLYRPSACPRCITIPSL 315

RESULT 13
US-09-880-503-5
; Sequence 5, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Root
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-5

Query Match      65.5%; Score 1508; DB 10; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.1e-124;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 KPSSPPEELKFCQGQKTLRPRFKIIGGFTTIIENQPFWFAAIYRRHSGSVTVVCGSLIS 195
Db 1 1 KPSSPPEELKFCQGQKTLRPRFKIIGGFTTIIENQPFWFAAIYRRHSGSVTVVCGSLIS 60
Qy 196 PCWVISATHCFIDYPKKEDIIVYLGRSRLNSNTQGMKEFEVENLILHKDYSADTLAHND 255
Db 61 PCWVISATHCFIDYPKKEDIIVYLGRSRLNSNTQGMKEFEVENLILHKDYSADTLAHND 120
Qy 256 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGKENSTDYLYPEQLKMT 315
Db 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGKENSTDYLYPEQLKMT 180
Qy 316 VVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTIGVS 375
Db 181 VVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTIGVS 240
Qy 376 WGRGCKALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
Db 241 WGRGCKALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276

RESULT 14
US-10-407-821-3
; Sequence 3, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRA-LEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
```



Db 1 IIGGEFTTIENQPMFAAIIYRRHGGSVTYVCGGSLMSPCWVISATHCFIDIPYKEDIVY 60  
QY 219 LGRSRLNSNTQGMKFEVENLILHKDYASDTLAHNDIALIKIRSKGRCQAQPSRTIQT 278  
Db 61 LGRSRLNSNTQGMKFEVENLILHKDYASDTLAHNDIALIKIRSKGRCQAQPSRTIQT 120  
QY 279 CLPSMYNDPQFGTSCEITGFKENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTK 338  
Db 121 CLPSMYNDPQFGTSCEITGFKENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTK 180  
QY 339 MLCAADPQWTKDSQCGSGGGLVCSLQGRMTLTGIVSGRGCAKDKPGVTVRVSHFLPW 398  
Db 181 MLCAADPQWTKDSQCGSGGGLVCSLQGRMTLTGIVSGRGCAKDKPGVTVRVSHFLPW 240  
QY 399 I 399  
Db 241 I 241

## RESULT 17

US-09-987-457-18  
; Sequence 18, Application US/09987457  
; Publication No. US20030013150A1  
; GENERAL INFORMATION:  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Manosroi, Jiradej  
; APPLICANT: Tayapiwatana, Chatchai  
; APPLICANT: Goetz, Friedrich  
; APPLICANT: Werner, Rolf-Guenther  
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes  
; FILE REFERENCE: 0652.218001  
; CURRENT APPLICATION NUMBER: US/09/987,457  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,573  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 00 27 782.2  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Homo sapiens (tPA)  
US-09-987-457-18

Query Match 37.6%; Score 864.5; DB 11; Length 527;  
Best Local Similarity 38.1%; Pred. No. 1.1e-67;  
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;  
QY 3 ELHQVP--SNCD---CLNGGTCVSNKYFSNIHWCNCPKFKGGOHCEIDKSKTCYEGNGHFY 58  
Db 42 QCHSVPVKSCSEPRCFNGGTCQQALYFSDP--VCQCPGFGAGKCCIEDTRATCYEDQGISY 100  
QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPCYVQVG 118  
Db 101 RGTWSTAESAECTNMNSALAQPYSGRRPDAILRLGLGNHNYCRNPNDRDSKPCWYVFKA 160  
QY 119 LKPLVQCMVHDCADG-----KGPS 138  
Db 161 GYSSEFCSTPACSEGNSDCYFNGSAYRGTHSLTESGASCLPWNSMILIGKYVTAQNP 220  
QY 139 -----SPPEELK-----FOCG-QKTLRPRFKIIGGE 163  
Db 221 AQALGLGKHNYCRNPDGAKPWCHLVKNRRLTWECYDVPSCSTGLRQYQPOQPRFKIGGL 280  
QY 164 FTTIENQPMFAAIIYRRH--RGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVLGRS 222  
Db 281 FADIASHPWQAIIFAKHRRSPGGERFLCGGILISSCWILSAAHCFQERFPPHLLTVILGRT 340  
QY 223 RLNSNTQGMKFEVENLILHKDYASDTLAHNDIALIKIRSKGRCQAQPSRTIQTICLPS 282  
Db 341 YRVVPGEEQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 398  
QY 283 RLNSNTQGMKFEVENLILHKDYASDTLAHNDIALIKIRSKGRCQAQPSRTIQTICLPS 282  
Db 341 YRVVPGEEQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 398

QY 283 MYNDPQFGTSCEITGFKENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCA 342  
Db 399 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCA 458  
QY 343 AD-----PQWKT--DSCQDGGGLVCSLQGRMTLTGIVSGRGCAKDKPGVTVRVSHFL 396  
Db 459 GDTSGGFPQANLHDACQDGGGLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYL 518  
QY 397 PWIRSHTK 404  
Db 519 DWIRDNR 526

## RESULT 18

US-09-987-455-19  
; Sequence 19, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.219001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-987-455-19

Query Match 37.6%; Score 864.5; DB 11; Length 527;  
Best Local Similarity 38.1%; Pred. No. 1.1e-67;  
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;  
QY 3 ELHQVP--SNCD---CLNGGTCVSNKYFSNIHWCNCPKFKGGOHCEIDKSKTCYEGNGHFY 58  
Db 42 QCHSVPVKSCSEPRCFNGGTCQQALYFSDP--VCQCPGFGAGKCCIEDTRATCYEDQGISY 100  
QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPCYVQVG 118  
Db 101 RGTWSTAESAECTNMNSALAQPYSGRRPDAILRLGLGNHNYCRNPNDRDSKPCWYVFKA 160  
QY 119 LKPLVQCMVHDCADG-----KGPS 138  
Db 161 GYSSEFCSTPACSEGNSDCYFNGSAYRGTHSLTESGASCLPWNSMILIGKYVTAQNP 220  
QY 139 -----SPPEELK-----FOCG-QKTLRPRFKIIGGE 163  
Db 221 AQALGLGKHNYCRNPDGAKPWCHLVKNRRLTWECYDVPSCSTGLRQYQPOQPRFKIGGL 280  
QY 164 FTTIENQPMFAAIIYRRH--RGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVLGRS 222  
Db 281 FADIASHPWQAIIFAKHRRSPGGERFLCGGILISSCWILSAAHCFQERFPPHLLTVILGRT 340  
QY 223 RLNSNTQGMKFEVENLILHKDYASDTLAHNDIALIKIRSKGRCQAQPSRTIQTICLPS 282  
Db 341 YRVVPGEEQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 398  
QY 283 MYNDPQFGTSCEITGFKENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCA 342  
Db 399 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCA 458

```
QY 343 AD-----POWKT-DSQGDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 396
Db 459 GDRSGGPQANLHDA CQDGGGGLVCLNDGRMTLVGLIISWGLGCGQKDPGVYTKVTNYL 518

QY 397 PWIRSHK 404
Db 519 DWIRDNMR 526

RESULT 19
US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-271-7

Query Match 37.6%; Score 864.5; DB 9; Length 562;
Best Local Similarity 38.1%; Pred. No. 1.2e-67;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;

QY 3 ELHQVP-SNCD----CLNGGTCVSNKYFNSIHWNCPCPKFGGQHCIEDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSEPRCFCNGGTCQOALYFSDP-VQCQPEGFAGKCCIEDTRATCYEDQGISY 135

QY 59 RGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPDRRRRCWCVQVG 118
Db 136 RGTWSTAESAETNNSSALAKPYSGRRPDAIRLGLGNHNYCRNPDRDRSKPWCYVFKA 195

QY 119 LKPLVQECMVHDCADG-----KPKS 138
Db 196 GKYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLRWNSMILIGKVYTAQNP 255

QY 139 -----SPPEELK-----FOCG-QKTLRPRFKIIGGE 163
Db 256 AQALGLGKKNYCRNPQGDAPKPVCHLVKNRRLTWECYCDVPSCSTCGLRQYSQPOFRKXGGL 315

QY 164 FTTIENQPFAPAIYRRH-RGGSVTVYVCGGSLISPCWVISATHCFIDYPPKEDYIVVLGRS 222
Db 316 FADIASHPPQAAIFAKHRRSPGERFLCGGILISSCWILSAHCFQERFPFPHHLTVILGRT 375

QY 223 RLNSNTQGMKEVENLILHKDYSATLAHNDIALKIRSKBGRCAQPSRTIQTICLPS 282
Db 376 YRVVPGEERQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433

QY 283 MYNDPQFGNSTCEITGFGKENSIDYLPQOLKMTVVKLI SHRECQPHYYGSEVTTKMLCA 342
Db 434 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHLVLPSSRCTSQHLLNRVTVDNMLCA 493

QY 343 AD-----POWKT-DSQGDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 396
Db 494 GDRSGGPQANLHDA CQDGGGGLVCLNDGRMTLVGLIISWGLGCGQKDPGVYTKVTNYL 553

QY 397 PWIRSHK 404
Db 554 DWIRDNMR 561
```

RESULT 20  
US-09-974-298-145

```
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 145
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145
```

```
Query Match 37.6%; Score 864.5; DB 10; Length 562;
Best Local Similarity 38.1%; Pred. No. 1.2e-67;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;

QY 3 ELHQVP-SNCD----CLNGGTCVSNKYFNSIHWNCPCPKFGGQHCIEDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSEPRCFCNGGTCQOALYFSDP-VQCQPEGFAGKCCIEDTRATCYEDQGISY 135

QY 59 RGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPDRRRRCWCVQVG 118
Db 136 RGTWSTAESAETNNSSALAKPYSGRRPDAIRLGLGNHNYCRNPDRDRSKPWCYVFKA 195

QY 119 LKPLVQECMVHDCADG-----KPKS 138
Db 196 GKYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLRWNSMILIGKVYTAQNP 255

QY 139 -----SPPEELK-----FOCG-QKTLRPRFKIIGGE 163
Db 256 AQALGLGKKNYCRNPQGDAPKPVCHLVKNRRLTWECYCDVPSCSTCGLRQYSQPOFRKXGGL 315

QY 164 FTTIENQPFAPAIYRRH-RGGSVTVYVCGGSLISPCWVISATHCFIDYPPKEDYIVVLGRS 222
Db 316 FADIASHPPQAAIFAKHRRSPGERFLCGGILISSCWILSAHCFQERFPFPHHLTVILGRT 375

QY 223 RLNSNTQGMKEVENLILHKDYSATLAHNDIALKIRSKBGRCAQPSRTIQTICLPS 282
Db 376 YRVVPGEERQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433

QY 283 MYNDPQFGNSTCEITGFGKENSIDYLPQOLKMTVVKLI SHRECQPHYYGSEVTTKMLCA 342
Db 434 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHLVLPSSRCTSQHLLNRVTVDNMLCA 493

QY 343 AD-----POWKT-DSQGDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 396
Db 494 GDRSGGPQANLHDA CQDGGGGLVCLNDGRMTLVGLIISWGLGCGQKDPGVYTKVTNYL 553

QY 397 PWIRSHK 404
Db 554 DWIRDNMR 561
```

RESULT 21  
US-10-443-701-4  
; Sequence 4, Application US/10443701  
; Publication No. US20030199016A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Yuan  
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS  
; FILE REFERENCE: P1788R1  
; CURRENT APPLICATION NUMBER: US/10/443,701  
; CURRENT FILING DATE: 2003-05-21  
; PRIOR APPLICATION NUMBER: US/09/703,695

```
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-443-701-4

Query Match      37.6%; Score 864.5; DB 12; Length 562;
Best Local Similarity 38.1%; Pred. No. 1.2e-67;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFNSIHCNCPKFGQCEIDKSKTCYEGNGHFY 58
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
77 QCHSVFVKSCSEPCFNGGTCQALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 135
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
59 RGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVG 118
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
136 RGTWSTAESGAECTNNSSALAOQPYSGRRPDALRLGLGNHNYCRNPDRDRSKPCYVYFKA 195
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
119 LKPLVQECMVHDCADG-----KKPS 138
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
196 GKYSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWN SMILIGKVYTAQNPS 255
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
139 -----SPPEELK-----FOCG-QKTLRPRFKIIGGE 163
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
256 AQALGLGKHNYCRNPDCGDAKPWCHVLKNRRLTWECYDVPSCTCGLRQYSPQPRFKIGGL 315
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
164 FTIENQPFALYRRH-RGGSVTVYVCGGSLISPCWVISATHCFIDYPPKEDIVYVIGRS 222
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
316 FADIASHPWQAALFAKHRSRSPGERFLCGGILISSCWILSAHCFQERFPFHLLTVILGRT 375
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
223 RLNSNTQGMKFVENLILHKDYSADTLAHNDIALKIRSKGRCAQPSRTIQTICLPS 282
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
376 YRVVPGEEOQKFEVKYIVHKEPDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
283 MYNDPQGTSCETGFGKENS TDLYPEQLKMTVVVKKLI SHRECOQPHYGVSEVTTKMLCA 342
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
434 ADLQLPDWTCELSGYGKHALSPFYSERLKEAHRVLYPSSRCTSQHLLNRTVTDNMLCA 493
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
343 AD-----POWKT--DSCQDGGPLVCSLQGRMTLTGIVSWGRCALCKDQGVYTRVSHFL 396
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
494 GDTSGGPGQANLHDACQDSDGGPLVCLNDGRMTLVGLISWGLGCGQKDVFGVYTKVTNYL 553
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
397 PWIRSHTK 404
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
554 DWIRDNMR 561

RESULT 23
US-09-880-503-8
; Sequence 8, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Rouf
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 8
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-8

Query Match      36.4%; Score 837; DB 10; Length 143;
Best Local Similarity 100.0%; Pred. No. 5.8e-66;
; LENGTH: 562
```

```
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNKCNPNRRPWCYVQVGLK 120
Db 61 KASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNKCNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPPSPPEE 143
Db 121 PLVQECMVHDCADGKPPSPPEE 143
RESULT 24
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4
Query Match 34.5%; Score 793; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 4e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNKCNPNRRPWCYVQVGLK 120
Db 61 KASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNKCNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135
RESULT 25
US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitten, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
```

```
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12
Query Match 34.5%; Score 793; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.1e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNKCNPNRRPWCYVQVGLK 120
Db 64 KASTDTMGPRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNKCNPNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138
RESULT 26
US-10-237-667-12
; Sequence 12, Application US/10237667
; Publication No. US20030022308A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitten, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
```





```
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: Macintosh
/ OPERATING SYSTEM: System 7.1
/ SOFTWARE: Word 5.1 (Patentin)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/237,866
/ FILING DATE: 10-Sep-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/797,689
/ FILING DATE: 31-JAN-1997
/ APPLICATION NUMBER: US 08/256,927
/ FILING DATE: 28-JUL-1994
/ APPLICATION NUMBER: FR 92/01064
/ FILING DATE: 31-JAN-1992
/ APPLICATION NUMBER: PCI/FR93/00085
/ FILING DATE: 28-JAN-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Smith Ph.D., Julie K.
/ REGISTRATION NUMBER: P-38,619
/ REFERENCE/DOCKET NUMBER: ST92006-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (610) 454-3839
/ TELEFAX: (610) 454-3808
/ INFORMATION FOR SEQ ID NO: 12:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 138 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match 34.5%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.1e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFGYRG 60
Db 4 SNELHVPNSCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFGYRG 63
QY 61 KASTDTMGPRCLPWSATVLQQTYHAHRSDALQGLGKHNKNCNPNRRRPWCYVQVGLK 120
Db 64 KASTDTMGPRCLPWSATVLQQTYHAHRSDALQGLGKHNKNCNPNRRRPWCYVQVGLK 123
QY 121 PLVQECWVHDCADGK 135
Db 124 PLVQECWVHDCADGK 138

RESULT 29
US-10-237-871-12
/ Sequence 12, Application US/10237871
/ Publication No. US20030036172A1
/ GENERAL INFORMATION:
/ APPLICANT: Fleer, Reinhard
/ Fournier, Alain
/ Guitton, Jean-Dominique
/ Jung, Gerard
/ Yeh, Patrice
/ TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
/ PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
/ CONTAINING SAID POLYPEPTIDES
/ NUMBER OF SEQUENCES: 36
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Rhone-Poulenc Rorer Inc.
/ STREET: 500 Arcola Road, 3C43
/ CITY: Collegeville
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19426
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/

MEDIUM TYPE: Floppy disk
/ COMPUTER: Macintosh
/ OPERATING SYSTEM: System 7.1
/ SOFTWARE: Word 5.1 (Patentin)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/237,871
/ FILING DATE: 10-Sep-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/797,689
/ FILING DATE: 31-JAN-1997
/ APPLICATION NUMBER: US 08/256,927
/ FILING DATE: 28-JUL-1994
/ APPLICATION NUMBER: FR 92/01064
/ FILING DATE: 31-JAN-1992
/ APPLICATION NUMBER: PCI/FR93/00085
/ FILING DATE: 28-JAN-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Smith Ph.D., Julie K.
/ REGISTRATION NUMBER: P-38,619
/ REFERENCE/DOCKET NUMBER: ST92006-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (610) 454-3839
/ TELEFAX: (610) 454-3808
/ INFORMATION FOR SEQ ID NO: 12:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 138 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-871-12

Query Match 34.5%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.1e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFGYRG 60
Db 4 SNELHVPNSCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFGYRG 63
QY 61 KASTDTMGPRCLPWSATVLQQTYHAHRSDALQGLGKHNKNCNPNRRRPWCYVQVGLK 120
Db 64 KASTDTMGPRCLPWSATVLQQTYHAHRSDALQGLGKHNKNCNPNRRRPWCYVQVGLK 123
QY 121 PLVQECWVHDCADGK 135
Db 124 PLVQECWVHDCADGK 138

RESULT 30
US-10-237-624-12
/ Sequence 12, Application US/10237624
/ Publication No. US20030082747A1
/ GENERAL INFORMATION:
/ APPLICANT: Fleer, Reinhard
/ Fournier, Alain
/ Guitton, Jean-Dominique
/ Jung, Gerard
/ Yeh, Patrice
/ TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
/ PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
/ CONTAINING SAID POLYPEPTIDES
/ NUMBER OF SEQUENCES: 36
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Rhone-Poulenc Rorer Inc.
/ STREET: 500 Arcola Road, 3C43
/ CITY: Collegeville
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19426
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: Macintosh
```

OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (Patentin)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,624  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-237-624-12

Query Match 34.5%; Score 793; DB 15; Length 138;  
Best Local Similarity 100.0%; Pred. No. 4.1e-62;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 SNELHQVPSNCDCLNGTCVSNKYFSNIHMCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
4 SNELHQVPSNCDCLNGTCVSNKYFSNIHMCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 63  
Qy 61 KASDTWTGRPCLPWNSATVLOOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
64 KASDTWTGRPCLPWNSATVLOOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 123  
Qy 121 PLVQECMVHDCADGK 135  
Db ||||||||||||||||  
124 PLVQECMVHDCADGK 138

Search completed: December 3, 2003, 15:05:43  
Job time : 44.1951 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:35:43 ; Search time 6.98276 Seconds  
(without alignments)  
818.010 Million cell updates/sec

Title: US-09-880-503-4

Perfect score: 793

Sequence: 1 SNEHLQVPSNCDCINGTCV.....QVGLKPLVQECMVHDCADGK 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	793	100.0	138	2	US-08-797-689-12
2	793	100.0	200	4	US-09-101-272G-73
3	793	100.0	208	4	US-09-101-272G-98
4	793	100.0	411	1	US-08-087-163-1
5	793	100.0	411	1	US-08-286-748B-18
6	793	100.0	411	1	US-08-153-799-18
7	793	100.0	430	1	US-07-942-157A-3
8	793	100.0	430	6	5219569-2
9	793	100.0	431	4	US-08-101-272G-1
10	793	100.0	431	6	518829-1
11	788	99.4	194	4	US-09-101-272G-80
12	788	99.4	201	3	US-09-101-272G-96
13	783	98.7	411	3	US-09-181-816-1
14	782	98.6	411	2	US-08-560-098A-48
15	769	97.0	157	3	US-08-142-590B-25
16	515	64.9	432	2	US-08-560-098A-47
17	514	64.8	365	1	US-08-093-741-83
18	514	64.8	365	1	US-08-720-012-83
19	514	64.8	393	2	US-08-560-098A-44
20	514	64.8	393	3	US-08-967-024C-24
21	514	64.8	393	3	US-08-967-024C-25
22	513	64.7	89	4	US-09-101-272G-62
23	335.5	42.3	477	2	US-08-560-098A-51
24	328.5	41.4	527	1	US-07-609-510B-16
25	328.5	41.4	527	2	US-08-811-949-39
26	328.5	41.4	527	5	PCT-US91-01025A-2
27	328.5	41.4	527	6	5185259-8

28	328.5	41.4	527	6	5520913-1
29	328.5	41.4	546	6	5200340-6
30	328.5	41.4	562	2	US-08-811-949-43
31	328.5	41.4	562	2	US-08-560-098A-50
32	328.5	41.4	562	2	US-08-883-795A-38
33	328.5	41.4	562	6	5185259-3
34	328.5	41.4	562	6	5200340-2
35	328.5	41.4	562	6	534773-2
36	328.5	41.4	562	6	524476-5
37	300	37.8	233	3	US-08-438-745-15
38	300	37.8	233	3	US-08-438-745-17
39	300	37.8	233	3	US-09-219-019-15
40	300	37.8	233	3	US-09-219-019-17
41	300	37.8	233	5	PCT-US94-05669A-15
42	300	37.8	233	5	PCT-US94-05669A-17
43	300	37.8	235	3	US-08-438-745-13
44	300	37.8	235	3	US-09-219-019-13
45	300	37.8	235	5	PCT-US94-05669A-13

## ALIGNMENTS

### RESULT 1

US-08-797-689-12  
; Sequence 12, Application US/08797689  
; Patent No. 5876569  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; APPLICANT: Fournier, Alain  
; APPLICANT: Guittion, Jean-Dominique  
; APPLICANT: Jung, Gerard  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D., Julie K.  
; REGISTRATION NUMBER: P-38,619  
; REFERENCE/DOCKET NUMBER: ST92006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 138 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

```
; MOLECULE TYPE: protein
US-08-797-689-12

Query Match      100.0%; Score 793; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.9e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 63

Qy 61 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 64 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 123

Qy 121 PLVOECMVHDCADGK 135
Db 124 PLVOECMVHDCADGK 138

RESULT 2
US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101.272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: ATF domain of uPA
US-09-101-272G-73

Query Match      100.0%; Score 793; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 4.4e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80

Qy 61 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 81 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140

Qy 121 PLVOECMVHDCADGK 135
Db 141 PLVOECMVHDCADGK 155

RESULT 3
US-09-101-272G-98
; Sequence 98, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101.272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFHI-ML chimeric protein
US-09-101-272G-98

Query Match      100.0%; Score 793; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 4.6e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 2 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 61

Qy 61 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 62 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 121

Qy 121 PLVOECMVHDCADGK 135
Db 122 PLVOECMVHDCADGK 136

RESULT 4
US-08-087-163-1
; Sequence 1, Application US/08087163
; Patent No. 5472692
; GENERAL INFORMATION:
; APPLICANT: Liu, Jian-Ning
; APPLICANT: Gurewich, Victor
; TITLE OF INVENTION: PRO-UKONINASE MUTANTS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,163
; FILING DATE: 07/02/93
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04353/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
US-08-087-163-1

Query Match      100.0%; Score 793; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 9.9e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
DB J SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRCLPWSATVLQOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
DB 61 KASTDTMGRCLPWSATVLQOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135  
RESULT 5  
US-08-286-748B-18  
; Sequence 18, Application US/08286748B  
; Patent No. 5759542  
; GENERAL INFORMATION:  
; APPLICANT: Victor Gurewich  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY  
; OF DRUGS BY PLATELETS FOR THE TREATMENT OF  
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3 5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 50Z or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: Wordperfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,748B  
; FILING DATE: August 5, 1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: J. Peter Fasse  
; REGISTRATION NUMBER: 32,983  
; REFERENCE/DOCKET NUMBER: 04547/013001  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 411  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-286-748B-18  
Query Match 100.0%; Score 793; DB 1; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.9e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRCLPWSATVLQOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
DB 61 KASTDTMGRCLPWSATVLQOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135

DB 121 PLVQECMVHDCADGK 135  
RESULT 6  
US-08-153-799-18  
; Sequence 18, Application US/08153799  
; Patent No. 5766883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: R Hain Swepe, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07974  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB90/00650  
; FILING DATE: 26-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/775952  
; FILING DATE: 29-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Swepe, R Hain  
; REGISTRATION NUMBER: 24864  
; REFERENCE/DOCKET NUMBER: 92H832  
; TELEPHONE: (908) 665 2400  
; TELEFAX: (908) 771 6159  
; TELEX: 219484  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 411 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-153-799-18  
Query Match 100.0%; Score 793; DB 1; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.9e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRCLPWSATVLQOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
DB 61 KASTDTMGRCLPWSATVLQOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135

RESULT 7  
US-07-942-157A-3  
; Sequence 3, Application US/07942157A  
; Patent No. 5648253  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Cha-Mer  
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/942,157A  
; FILING DATE: 19920908  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/631673  
; FILING DATE: 20-DEC-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: TS1108Cont.  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404)815-6508  
; TELEFAX: (404)815-6555  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 430 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEetical: NO  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..19  
; OTHER INFORMATION: /label= peptide  
; OTHER INFORMATION: /note= "MAP signal"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 198..203  
; OTHER INFORMATION: /label= modified  
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"  
US-07-942-157A-3

Query Match 100.0%; Score 793; DB 1; Length 430;  
Best Local Similarity 100.0%; Pred. No. 1e-71;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60  
Db 20 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 79  
QY 61 KASDTDMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVOVGLK 120  
Db 80 KASDTDMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVOVGLK 139  
QY 121 PLVQECMVHDCADGK 135  
Db 140 PLVQECMVHDCADGK 154

RESULT 8  
5219569-2  
; Patent No. 5219569

; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR, GORDON A.  
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE  
; NUMBER OF SEQUENCES: 6  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/766,858  
; FILING DATE: 16-AUG-1985  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 725,468  
; FILING DATE: 22-APR-1985  
; SEQ ID NO: 2:  
; LENGTH: 430  
; 5219569-2  
Query Match 100.0%; Score 793; DB 6; Length 430;  
Best Local Similarity 100.0%; Pred. No. 1e-71;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60  
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80  
QY 61 KASDTDMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVOVGLK 120  
Db 81 KASDTDMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVOVGLK 140  
QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155  
RESULT 9  
US-09-101-272G-1  
; Sequence 1, Application US/09101272G  
; Patent No. 6509445  
; GENERAL INFORMATION:  
; APPLICANT: Nissin Food Products Co., Ltd.  
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
; FILE REFERENCE: Q50979  
; CURRENT APPLICATION NUMBER: US/09/101,272G  
; CURRENT FILING DATE: 1998-07-08  
; PRIOR APPLICATION NUMBER: JP 1059/1996  
; PRIOR FILING DATE: 1996-01-08  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1:  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (21)..()  
; OTHER INFORMATION:  
; NAME/KEY: misc feature  
; LOCATION: (20)..()  
; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)  
US-09-101-272G-1

Query Match 100.0%; Score 793; DB 4; Length 431;  
Best Local Similarity 100.0%; Pred. No. 1e-71;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60  
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80  
QY 61 KASDTDMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVOVGLK 120  
Db 81 KASDTDMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVOVGLK 140  
QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155

```
RESULT 10
5188829-1
; Patent No. 5188829
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIKAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1988
; SEQ ID NO: 1:
; LENGTH: 431
5188829-1
Query Match 100.0%; Score 793; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 1e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
Db 21 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 80
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
Db 141 PLVQECMVHDCADGK 155
RESULT 11
US-09-101-272G-80
; Sequence 80, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: O50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 80:
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATPHI chimeric protein
US-09-101-272G-80
Query Match 99.4%; Score 788; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
Db 2 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 61
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 62 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 121
QY 121 PLVQECMVHDCADG 134
Db 122 PLVQECMVHDCADG 135
RESULT 12
US-09-101-272G-96
; Sequence 96, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: O50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 96:
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATPHI-CL chimeric protein
US-09-101-272G-96
Query Match 99.4%; Score 788; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 1.4e-71;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
Db 2 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 61
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 62 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 121
QY 121 PLVQECMVHDCADG 134
Db 122 PLVQECMVHDCADG 135
RESULT 13
US-09-181-816-1
; Sequence 1, Application US/09181816
; Patent No. 6277818
; GENERAL INFORMATION:
; APPLICANT: JONES, Andrew P.
; APPLICANT: JONES, Terence R.
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
; TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
; FILE REFERENCE: 329042000300 SIDN 1-7
; CURRENT APPLICATION NUMBER: US/09/181,816
; CURRENT FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1:
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-181-816-1
Query Match 98.7%; Score 783; DB 3; Length 411;
Best Local Similarity 99.3%; Pred. No. 1e-70;
Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
Db 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135
```



```
RESULT 14
US-08-560-098A-48
; Sequence 48, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-48
Query Match 98.6%; Score 782; DB 2; Length 411;
Best Local Similarity 99.3%; Pred. No. 1.3e-70;
Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGTCTVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFG 60
Db 1 SNELHQPNSCDCLNGTCTVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFG 60

QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNCRPNRRPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNCRPNRRPWCYVQVGLK 120

QY 121 PLVQECWVHDCADGK 135
Db 121 PLVQECWVHDCADGK 135

RESULT 15
US-08-142-590B-25
; Sequence 25, Application US/08142590B
; Patent No. 6120765
; GENERAL INFORMATION:
; APPLICANT: HIBINO, Tashihiko, Takahashi, Tadahito, Hori, Izumi, and
; TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,590B
; FILING DATE: 25-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,318
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: MGP-009CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-142-590B-25
Query Match 97.0%; Score 769; DB 3; Length 157;
Best Local Similarity 97.0%; Pred. No. 8.5e-70;
Matches 131; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGTCTVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFG 60
Db 1 SNELHQPNSCDCLNGTCTVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFG 60

QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNCRPNRRPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNCRPNRRPWCYVQVGLK 120

QY 121 PLVQECWVHDCADGK 135
Db 121 PLVQECWVHDCADGK 135

RESULT 16
US-08-560-098A-47
; Sequence 47, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA: P 44 40 892.7
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-47

Query Match 64.9%; Score 515; DB 2; Length 432;
Best Local Similarity 75.8%; Pred. No. 7.6e-44;
Matches 93; Conservative 8; Mismatches 16; Indels 6; Gaps 1;

QY 13 CLNGTCVSNKYPFNIHWCNCPKFGQHCQCEIDKSKTCYEGNGHFGYRGKASTDTMGRPCL 72
Db 40 CVTGEPTKPESHNNGDFEIPREY-----LQISKTCYEGNGHFGYRGKASTDTMGRPCL 93
QY 73 PWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMVHDC 132
Db 94 PWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMVHDC 153
QY 133 DGK 135
Db 154 DGK 156

RESULT 17
US-08-093-741-83
; Sequence 83, Application US/08093741
; Patent No. 5681721
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, Gerd J.
; APPLICANT: WENNDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: SAUNDERS, DEREK J.
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N. W. Suite 700
; CITY: Washington, D.C.
; COUNTRY: U.S.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 20-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/093,741
; FILING DATE: 20-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-720-012-83

Query Match 64.8%; Score 514; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 7.9e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFGYRGKASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPD 106
Db 1 SKTCYEGNGHFGYRGKASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPD 60
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 61 NRRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 18
US-08-720-012-83
; Sequence 83, Application US/08720012
; Patent No. 5747291
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, Gerd J.
; APPLICANT: WENNDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: SAUNDERS, DEREK J.
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N. W. Suite 700
; CITY: Washington, D.C.
; COUNTRY: U.S.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/093,741
; FILING DATE: 20-JUL-1993
; APPLICATION NUMBER: DE P43 23 754.1
; FILING DATE: 15-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-720-012-83

Query Match 64.8%; Score 514; DB 1; Length 365;
```

```
Best Local Similarity 100.0%; Pred. No. 7.9e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPD 106
Db 1 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPD 60
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 61 NRRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 19
US-08-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEBDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSCHA, Elke
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 44:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: protein
; MOLECULE TYPE: protein
US-08-560-098A-44

Query Match 64.8%; Score 514; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 8.6e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPD 61
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGK 90

RESULT 20
US-08-967-024C-24
; Sequence 24, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNEBDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSCHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 665.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: protein
; MOLECULE TYPE: protein
US-08-967-024C-24

Query Match 64.8%; Score 514; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 8.6e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPD 61
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGK 90

RESULT 21
US-08-967-024C-25
; Sequence 25, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNEBDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSCHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/967,024C  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: P 44 42 665.8  
FILING DATE: 30-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/42444  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-8800  
TELEFAX: (202) 628-8844  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 393 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-967-024C-25

Query Match 64.8%; Score 514; DB 3; Length 393;  
Best Local Similarity 100.0%; Pred. No. 8.6e-44;  
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 SKTCYEGNGHYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPD 106

Db 2 SKTCYEGNGHYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPD 61

Qy 107 NRRPWCYVQVGLKPLVQECMWHDCADGX 135

Db 62 NRRPWCYVQVGLKPLVQECMWHDCADGX 90

## RESULT 22

US-09-101-272G-62  
Sequence 62, Application US/09101272G  
Patent No. 6509445  
GENERAL INFORMATION:  
APPLICANT: Nissin Food Products Co., Ltd.  
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
FILE REFERENCE: Q50979  
CURRENT APPLICATION NUMBER: US/09/101,272G  
CURRENT FILING DATE: 1998-07-08  
PRIOR APPLICATION NUMBER: JP 1059/1996  
PRIOR FILING DATE: 1996-01-08  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 62  
LENGTH: 89  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: residues 43-131 of the ATF domain of uPA  
US-09-101-272G-62

Query Match 64.7%; Score 513; DB 4; Length 89;  
Best Local Similarity 100.0%; Pred. No. 2e-44;  
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 EIDSKTCYEGNGHYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYC 102

Db 1 EIDSKTCYEGNGHYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYC 60

Qy 103 RNPDRRPPWCYVQVGLKPLVQECMWHDC 131

|||||

Db 61 RNPDRRPPWCYVQVGLKPLVQECMWHDC 89

## RESULT 23

US-08-560-098A-51  
Sequence 51, Application US/08560098A  
Patent No. 5976841  
GENERAL INFORMATION:  
APPLICANT: WNEIDT, Stephan  
APPLICANT: HEINZEL-WIELAND, Regina  
APPLICANT: STEFFENS, Gerd Josef  
TITLE OF INVENTION: Proteins having Fibrinolytic and  
TITLE OF INVENTION: Coagulation-inhibiting Properties  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
STREET: 1200 G Street, N.W., Suite 700  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/560,098A  
FILING DATE: 17-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: P 44 40 892.7  
FILING DATE: 17-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/42448  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-8800  
TELEFAX: (202) 628-8844  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 477 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-560-098A-51

Query Match 42.3%; Score 335.5; DB 2; Length 477;  
Best Local Similarity 46.3%; Pred. No. 8.3e-26;  
Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

Qy 3 ELHQVP----SNCDCLNGTGVSNKYFSNIHWCNCPKFGGHCHEIDKSKTCYEGNGHYF 58

Db 78 QCHTVFKVSCSELRCFNGGTCAQAAASFSDF-VCCPKGTGKQCEVDTHATCYKQGVY 136

Qy 59 RKGASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVG 118

Db 137 RGTWSTSESQAQCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPNPNNSKPCWYVKA 196

Qy 119 LKPLVQECMWHDC 132

Db 197 SKFILEFCVPVCS 210

## RESULT 24

US-07-609-510B-16  
Sequence 16, Application US/07609510B  
Patent No. 5326700  
GENERAL INFORMATION:  
APPLICANT: Berg et al.  
TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue I  
NUMBER OF SEQUENCES: 16

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,510B
; FILING DATE: 19901106
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-609-510B-16

Query Match 41.4%; Score 328.5; DB 1; Length 527;
Best Local Similarity 46.3%; Pred. No. 4.7e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELQVPS-SNCD---CLNGGTCVSNKYFNSIHWNCNCPKFGGQHCIDKSKTCYEGNGHFY 58
Db 42 QCHSVFVKSCSEPCFNGGTCQALYFSDP-VQCPEGFAGKCCIEDTRATCYEDQGGSY 100
QY 59 RGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNCRPNDRRPPWCYVQVG 118
Db 101 RGTWSTAESGAECTNWNSSALAKPYSGRRPDAIRLGLGNHNYCRNPD RDSKPCYVFKA 160
QY 119 LKPLVQECMVHDCADG 134
Db 161 GKYSEFCSTPACSEG 176

RESULT 25
US-08-811-949-39
; Sequence 39, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C. JEFFERSON DAVIS HIGHWAY, SUITE 400
; STREET: 1755 S.
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-811-949-39

Query Match 41.4%; Score 328.5; DB 2; Length 527;
Best Local Similarity 46.3%; Pred. No. 4.7e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELQVPS-SNCD---CLNGGTCVSNKYFNSIHWNCNCPKFGGQHCIDKSKTCYEGNGHFY 58
Db 42 QCHSVFVKSCSEPCFNGGTCQALYFSDP-VQCPEGFAGKCCIEDTRATCYEDQGGSY 100
QY 59 RGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNCRPNDRRPPWCYVQVG 118
Db 101 RGTWSTAESGAECTNWNSSALAKPYSGRRPDAIRLGLGNHNYCRNPD RDSKPCYVFKA 160
QY 119 LKPLVQECMVHDCADG 134
Db 161 GKYSEFCSTPACSEG 176

RESULT 26
PCT-US91-01025A-2
; Sequence 2, Application PC/TUS9101025A
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01025A
; FILING DATE: 19910214
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/486,657
; FILING DATE: 1 March 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet B.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 454P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; PCT-US91-01025A-2
```



```

; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 562 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-811-949-43

```

```

Query Match      41.4%; Score 328.5; DB 2; Length 562;
Best Local Similarity 46.3%; Pred. No. 5.1e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY      3 ELHQVP-SNCD---CLNGGTCVSNKYESNIHWCNCPKKGQGHCEIDKSKTCYEGNGHFY 58
Db      77 QCHSVPVKSCSEPRCFNGGTCQALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 135

QY      59 RGRKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVG 118
Db      136 RGTWSTAESGAECTNNSSALAKPYSGRPPDAIRLGLGNHNYCRNPDRDRDSKPCYVFKA 195

QY      119 LKPLVOECMVHDCADG 134
Db      196 GKYSEFCSTPACSEG 211

```

Search completed: December 3, 2003, 14:45:35  
Job time : 7.98276 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:02 ; Search time 21.2657 Seconds  
(without alignments)  
1007.637 Million cell updates/sec

Title: US-09-880-503-4  
Perfect score: 793  
Sequence: 1 SNELHQVPNCCLNGGTCV.....QVGLKPLVQCMVHDCADGK 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	793	100.0	135	AAE16545	Human urokinase-ty
2	793	100.0	143	AAE16549	Human uPA amino te
3	793	100.0	337	AAE16549	Human colon cancer
4	793	100.0	337	AAE16549	Human ovarian anti
5	793	100.0	403	AAE16547	Human urokinase-ty
6	793	100.0	411	AAE16547	Sequence encoded b
7	793	100.0	411	AAE16547	Human native prour
8	793	100.0	411	AAE16547	UK-S3 as encoded b
9	793	100.0	411	AAE16547	Urokinase precursor

10	793	100.0	411	12	AAE10057	Pro-urokinase deri
11	793	100.0	411	12	AAE10058	Pro-urokinase deri
12	793	100.0	411	14	AAE40225	PUK. Homo sapiens
13	793	100.0	411	16	AAE62991	Pro-urokinase. Ho
14	793	100.0	411	16	AAE62992	Pro-urokinase muta
15	793	100.0	411	16	AAE62993	Pro-urokinase muta
16	793	100.0	411	16	AAE62994	Pro-urokinase muta
17	793	100.0	411	16	AAE62995	Pro-urokinase muta
18	793	100.0	411	16	AAE62996	Pro-urokinase muta
19	793	100.0	411	16	AAE62997	Pro-urokinase muta
20	793	100.0	411	16	AAE62998	Pro-urokinase muta
21	793	100.0	411	16	AAE62999	Pro-urokinase muta
22	793	100.0	411	16	AAE63000	Pro-urokinase muta
23	793	100.0	411	16	AAE63001	Pro-urokinase muta
24	793	100.0	411	16	AAE63002	Pro-urokinase muta
25	793	100.0	411	16	AAE63003	Pro-urokinase muta
26	793	100.0	411	16	AAE63004	Pro-urokinase muta
27	793	100.0	411	16	AAE63005	Pro-urokinase muta
28	793	100.0	411	16	AAE63006	Pro-urokinase muta
29	793	100.0	411	16	AAE63007	Pro-urokinase muta
30	793	100.0	411	16	AAE63008	Pro-urokinase muta
31	793	100.0	411	17	AAE92926	Pro-urokinase. Ho
32	793	100.0	411	21	AAE92836	Urokinase plasmino
33	793	100.0	411	21	AAE16544	Human urokinase-ty
34	793	100.0	411	23	AAE16544	Recombinant single
35	793	100.0	424	12	AAE10334	Inhibitor resistan
36	793	100.0	430	18	AAE24578	Inhibitor resistan
37	793	100.0	431	6	AAE50114	Sequence encoded b
38	793	100.0	431	7	AAE60783	Human urokinase.
39	793	100.0	431	8	AAE70258	Sequence of human
40	793	100.0	431	8	AAE71451	Modified prourokin
41	793	100.0	431	8	AAE71663	Modified prourokin
42	793	100.0	431	8	AAE71698	Modified prourokin
43	793	100.0	431	8	AAE71699	Modified prourokin
44	793	100.0	431	9	AAE80430	Deduced AA sequenc
45	793	100.0	431	9	AAE81204	Pro-urokinase with

## ALIGNMENTS

RESULT 1  
AAE16545 standard; Protein; 135 AA.

AAE16545;  
09-APR-2002 (first entry)

Human urokinase-type plasminogen activator amino terminal fragment (ATF).  
Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
clotting disorder; uterine contraction disorder; respiratory disease;  
adult respiratory distress syndrome; amino terminal fragment; ATF;  
male impotence.

Homo sapiens.

WO200197752-A2.

27-DEC-2001.

13-JUN-2001; 2001WO-US18976.

20-JUN-2000; 2000US-212874P.

(UYFE-) UNIV PENNSYLVANIA.

Cines DB, Higazi AA;



DR WPI; 2002-122240/16.  
DR N-PSDB; AAD27078.  
XX  
PT Composition for modulating muscle cell and tissue contractility for  
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
PT comprising domains from urokinase-type plasminogen activator -  
XX  
XX Claim 11; Fig ID: 117pp; English.  
PS  
XX The invention relates to a composition comprising one or more domains of  
CC urokinase-type plasminogen activator (uPA). The composition is used to  
CC modulate the contractility and angiogenic activity of a mammalian muscle,  
CC endothelial cell or tissue. The composition is used for treating stroke,  
CC hypotension, hypertension, atherosclerosis, heart attack, microvascular  
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
CC diabetic retinopathy, wound healing, clotting disorder, uterine  
CC contraction disorder, male impotence, respiratory disease or condition  
CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
CC hypertension, microvascular thrombotic occlusion, and a disorder  
CC associated with chronic intrapulmonary fibrin formation. The present  
CC sequence is human urokinase-type plasminogen activator (uPA) amino  
CC terminal fragment (ATF).  
XX  
SQ Sequence 135 AA;  
Query Match 100.0%; Score 793; DB 23; Length 135;  
Best Local Similarity 100.0%; Pred. No. 7.4e-52;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Qy 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
Db 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
Qy 121 PLVQECMVHDCADGK 135  
Db 121 PLVQECMVHDCADGK 135  
RESULT 2  
AAE16549  
ID AAE16549 standard; Protein; 143 AA.  
AC AAE16549;  
DT 09-APR-2002 (first entry)  
XX  
XX Human uPA amino terminal fragment (ATF) and connecting peptide.  
DE  
XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
KW clotting disorder; uterine contraction disorder; respiratory disease;  
KW adult respiratory distress syndrome; amino terminal fragment; ATF;  
KW male impotence.  
XX  
XX Homo sapiens.  
XX  
XX WO2000197752-A2.  
XX  
XX 27-DEC-2001.  
XX  
XX 13-JUN-2001; 2001WO-US18976.  
XX  
XX 20-JUN-2000; 2000US-212874P.  
XX  
XX (UYPE-) UNIV PENNSYLVANIA.

XX Cines DB, Higazi AA;  
XX  
XX WPI; 2002-122240/16.  
DR N-PSDB; AAD27082.  
XX  
PT Composition for modulating muscle cell and tissue contractility for  
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
PT comprising domains from urokinase-type plasminogen activator -  
XX  
XX Claim 24; Fig 1H; 117pp; English.  
PS  
XX The invention relates to a composition comprising one or more domains of  
CC urokinase-type plasminogen activator (uPA). The composition is used to  
CC modulate the contractility and angiogenic activity of a mammalian muscle,  
CC endothelial cell or tissue. The composition is used for treating stroke,  
CC hypotension, hypertension, atherosclerosis, heart attack, microvascular  
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
CC diabetic retinopathy, wound healing, clotting disorder, uterine  
CC contraction disorder, male impotence, respiratory disease or condition  
CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
CC hypertension, microvascular thrombotic occlusion, and a disorder  
CC associated with chronic intrapulmonary fibrin formation. The present  
CC sequence is human urokinase-type plasminogen activator (uPA) amino  
CC terminal fragment (ATF) and connecting peptide.  
XX  
SQ Sequence 143 AA;  
Query Match 100.0%; Score 793; DB 23; Length 143;  
Best Local Similarity 100.0%; Pred. No. 7.8e-52;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Qy 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
Db 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
Qy 121 PLVQECMVHDCADGK 135  
Db 121 PLVQECMVHDCADGK 135  
RESULT 3  
AAG75492  
ID AAG75492 standard; Protein; 337 AA.  
AC AAG75492;  
DT 03-SEP-2001 (first entry)  
XX  
XX Human colon cancer antigen protein SEQ ID NO:6256.  
DE  
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma; chromosome 10.  
XX  
XX Homo sapiens.  
XX  
XX WO2000122920-A2.  
XX  
XX 05-APR-2001.  
XX  
XX 28-SEP-2000; 2000WO-US26524.  
XX  
XX 29-SEP-1999; 99US-0157137.  
PR 03-NOV-1999; 99US-0163280.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX WPI; 2001-235357/24.  
DR N-PSDB; AAH34897.  
XX  
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers -  
XX  
XX  
PS Claim 11; Page 7707-7708; 9803pp; English.  
XX  
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing  
CC inactive proteins or to supplement the patient's own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated Ps,  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
CC and AAH77789 represent sequences used in the exemplification of the  
CC present invention.  
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
XX  
XX  
SQ Sequence 337 AA;  
  
Query Match 100.0%; Score 793; DB 22; Length 337;  
Best Local Similarity 100.0%; Pred. No. 1.6e-51;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SNEHQVPSNCDCLNGTGVSNKYFNSHWCNCPKPKFGGHCIDKSKTCYEGNGHFYRG 60  
DB 27 SNEHQVPSNCDCLNGTGVSNKYFNSHWCNCPKPKFGGHCIDKSKTCYEGNGHFYRG 86  
  
QY 61 KASTDTMGRCPLPWSATVILQOTVHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
DB 87 KASTDTMGRCPLPWSATVILQOTVHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 146  
  
QY 121 PLVQECMVHDCADGK 135  
DB 147 PLVQECMVHDCADGK 161  
  
RESULT 4  
ABP41795  
XX  
XX AC ABP41795;  
XX  
XX 22-AUG-2002 (first entry)  
XX  
XX Human ovarian antigen HWCEB79, SEQ ID NO:2927.  
XX  
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
KW inflammatory condition; immune disorder; blood disorder;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disorder; urinary system disorder; drug screening;  
KW gene therapy; chromosome mapping; forensic analysis;  
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
KW antiinflammatory; gynaecological; reproductive; chromosome 10q24.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200200677-A1.  
PN

XX  
PD  
XX  
XX  
XX 03-JAN-2002.  
XX  
XX 07-JUN-2001; 2001WO-US18569.  
XX  
XX 07-JUN-2000; 2000US-209467P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Birse CE, Rosen CA;  
XX  
XX WPI; 2002-147878/19.  
DR N-PSDB; ABQ54872.  
XX  
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
PT useful in the prevention, treatment and diagnosis of cancer (e.g.  
XX ovarian cancer), immune disorders, cardiovascular disorders and  
PT neurological diseases -  
XX  
XX Claim 11; SEQ ID No 2927; 2922pp; English.  
XX  
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
CC to the sequences of the invention. The invention additionally relates to  
CC recombinant vectors and host cells comprising human ovarian antigen  
CC polynucleotides, antibodies against human ovarian antigens, and the use  
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
CC treating, prognosing or preventing various ovary and/or breast-related  
CC disorders. Such conditions include ovarian cancer and breast cancer, and  
CC metastatic tumours of ovarian or breast origin, reproductive system  
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
CC vaginitis), immune disorders (e.g., congenital and acquired  
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
CC respiratory disorders, neurological disorders, gastrointestinal disorders  
CC and urinary system disorders. Ovarian antigen polypeptides and  
CC polynucleotides may also be used in screening for compounds which  
CC modulate ovarian antigen expression or activity. The polynucleotides may  
CC further be used for gene therapy, chromosome mapping, in the  
CC identification of individuals and in forensic analysis, and the  
CC polypeptides may be used as food additives or to prepare antibodies  
CC useful in disease diagnosis, drug targeting and phenotyping. The present  
CC sequence represents a human ovarian antigen of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 337 AA;  
  
Query Match 100.0%; Score 793; DB 23; Length 337;  
Best Local Similarity 100.0%; Pred. No. 1.6e-51;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SNEHQVPSNCDCLNGTGVSNKYFNSHWCNCPKPKFGGHCIDKSKTCYEGNGHFYRG 60  
DB 27 SNEHQVPSNCDCLNGTGVSNKYFNSHWCNCPKPKFGGHCIDKSKTCYEGNGHFYRG 86  
  
QY 61 KASTDTMGRCPLPWSATVILQOTVHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120  
DB 87 KASTDTMGRCPLPWSATVILQOTVHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 146  
  
QY 121 PLVQECMVHDCADGK 135  
DB 147 PLVQECMVHDCADGK 161  
  
RESULT 5  
AAE16547  
ID AAE16547 standard; Protein; 403 AA.

XX AAE16547;  
AC 09-APR-2002 (first entry)  
XX Human urokinase-type plasminogen activator scUPA delta136-143 mutant.  
XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
KW clotting disorder; uterine contraction disorder; respiratory disease;  
KW male impotence; adult respiratory distress syndrome; scUPA delta136-143;  
KW single chain urokinase; mutant; mutein.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX WO200197752-A2.  
XX 27-DEC-2001.  
XX 13-JUN-2001; 2001WO-US18976.  
XX 20-JUN-2000; 2000US-212874P.  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX Cines DB, Higazi AA;  
XX WPI; 2002-122240/16.  
XX N-PSDB; AAD27080.  
XX Composition for modulating muscle cell and tissue contractility for  
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
PT comprising domains from urokinase-type plasminogen activator -  
XX Claim 22; Fig 1F; 117pp; English.  
XX The invention relates to a composition comprising one or more domains of  
CC urokinase-type plasminogen activator (uPA). The composition is used to  
CC modulate the contractility and angiogenic activity of a mammalian muscle,  
CC endothelial cell or tissue. The composition is used for treating stroke,  
CC hypotension, hypertension, atherosclerosis, heart attack, microvascular  
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
CC diabetic retinopathy, wound healing, clotting disorder, uterine  
CC contraction disorder, male impotence, respiratory disease or condition  
CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
CC hypertension, microvascular thrombotic occlusion, and a disorder  
CC associated with chronic intrapulmonary fibrin formation. The present  
CC sequence is human urokinase-type plasminogen activator (uPA) single  
CC chain urokinase (scUPA) deletion mutant designated as scUPA delta136-143.  
XX  
SQ Sequence 403 AA;  
Query Match 100.0%; Score 793; DB 23; Length 403;  
Best Local Similarity 100.0%; Pred. No. 1.9e-51;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDCLNGTCTVSNKYFNSIHWNCNPKFGGQHCIDSKTCYEGNGHGYRG 60  
DB 1 SNELHQVPSNCDCLNGTCTVSNKYFNSIHWNCNPKFGGQHCIDSKTCYEGNGHGYRG 60  
QY 61 KASTDTMGRCLPWSATVLQOTYHAHRSDALQLGLGKHYCNPNDRRPPWCYVOVGLK 120  
DB 61 KASTDTMGRCLPWSATVLQOTYHAHRSDALQLGLGKHYCNPNDRRPPWCYVOVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135

RESULT 6  
AAP50871  
ID AAP50871 standard; protein; 411 AA.  
XX  
XX AAP50871;  
XX 30-NOV-1991 (first entry)  
DT Sequence encoded by cDNA sequence for human urokinase zymogen  
DE (Japanese Patent Application No.37119/84).  
DE Thrombolytic agent; plasminogen activator activity; fibrin affinity;  
KW enzyme.  
KW  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Cleavage-site 158..159  
FT /note= "potential cleavage site which generates  
FT the two-chain form from the zymogen"  
FT  
FT Disulfide-bond 50..131  
FT Disulfide-bond 71..113  
FT Disulfide-bond 102..126  
FT Disulfide-bond 148..279  
FT Disulfide-bond 189..205  
FT Disulfide-bond 197..268  
FT Disulfide-bond 231..362  
FT Disulfide-bond 325..341  
FT Disulfide-bond 352..380  
XX  
PN EP139447-A.  
XX  
PD 02-MAY-1985.  
XX  
XX 07-SEP-1984; 84EP-0306117.  
XX 17-OCT-1983; 83JP-0195051.  
PR 13-SEP-1983; 83JP-0170354.  
XX  
XX (GREG ) GREEN CROSS CORP.  
XX Kasai S, Arimura H, Mori K, Nishida M, Suyama T;  
XX WPI; 1985-106530/18.  
XX New urokinase zymogen - useful as thrombolytic agent  
PT  
PS Disclosure; Page 12; 30pp; English.  
XX  
XX Zymogen AAP50871 is the inactive precursor form of human urokinase.  
CC Urokinase zymogen is cleaved into the two-chain form composed of  
CC characteristic urokinase H (molecular wt. of 30,000) and L (molecular  
CC wt. of 20,000) chains when treated with catalytic amounts of plasmin.  
CC The patentors claim a new urokinase zymogen which has mol. wt. ca.  
CC 50,000, a single chain molecular structure, and selective affinity  
CC for fibrin. It is a thrombolytic agent which manifests its  
CC plasminogen activator activity on cleavage by proteolytic enzymes  
CC (e.g. plasmin) and has higher affinity for fibrin than known forms  
CC of urokinase.  
XX  
SQ Sequence 411 AA;  
Query Match 100.0%; Score 793; DB 6; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.9e-51;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDCLNGTCTVSNKYFNSIHWNCNPKFGGQHCIDSKTCYEGNGHGYRG 60  
DB 1 SNELHQVPSNCDCLNGTCTVSNKYFNSIHWNCNPKFGGQHCIDSKTCYEGNGHGYRG 60  
QY 61 KASTDTMGRCLPWSATVLQOTYHAHRSDALQLGLGKHYCNPNDRRPPWCYVOVGLK 120

Db 61 KASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
Db 121 PLVQECMVHDCADGK 135

RESULT 7  
AAW13634  
ID AAW13634 standard; Protein; 411 AA.  
XX AC AAW13634;  
XX DT 04-JUN-1997 (first entry)  
XX DE Human native prourokinase.  
XX KW Human; prourokinase; hPUK; variant; half-life: increase; EGF;  
KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
FH Domain 10..42  
FH /label= EGF domain  
FH /note= "in\_Glained variants, at least part of the  
FT EGF domain is deleted (see comments)"  
FT Region 10..19  
FT /label= first\_loop  
FT Region 20..31  
FT /label= second\_loop  
FT Region 33..42  
FT /label= third\_loop  
XX EP398361-A.  
XX 22-NOV-1990.  
XX 18-MAY-1990; 90EP-0109472.  
XX 22-FEB-1990; 90JP-0042020.  
XX 18-MAY-1989; 89JP-0126433.  
XX 03-JUL-1986; 86JP-0156936.  
XX 18-FEB-1987; 87JP-0036495.  
XX 18-MAY-1989; 89JP-0126434.  
XX (GREG ) GREEN CROSS CORP.  
XX Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;  
XX Morita M, Tanabe T;  
XX WPI; 1990-350146/47.  
XX N-PSDB; AAT61671.  
XX Human pro-urokinase variants - deficient in loop regions of  
XX epidermal growth factor, showing long blood half-life, as  
XX fibrinolytic agent  
XX Disclosure; Fig 1; 22pp; English.  
XX New variants of human prourokinase (hPUK) comprise a hPUK deficient  
XX in (i) at least part of the first loop region of the epidermal growth  
XX factor (EGF) domain; (ii) at least part of the first loop and at  
XX least part of the second loop; or (iii) at least part of the third  
XX loop. The hPUK variants show an increased blood half-life comparable  
XX to that of the whole EGF domain-deficient hPUK variant and urokinase  
XX while retaining the same properties as those of hPUK. They have  
XX potent thrombolytic activity and very little tendency to cause  
XX spontaneous bleeding. The present sequence is that of the wild-type  
XX hPUK protein, including the EGF domain.  
XX Sequence 411 AA;

Query Match 100.0%; Score 793; DB 11; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.9e-51;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFEYRG 60  
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFEYRG 60  
QY 61 KASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120  
Db 61 KASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
Db 121 PLVQECMVHDCADGK 135

RESULT 8  
AAR05117  
ID AAR05117 standard; protein; 411 AA.  
XX AC AAR05117;  
XX DT 25-MAR-2003 (updated)  
DT 04-OCT-1990 (first entry)  
XX UK-S3 as encoded by PUKS3.  
XX Urokinase; glycosylation.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH misc\_difference 153  
FT /label=synthetic\_mutation  
FT /note="old seq (Leu)"  
FT misc\_difference 155  
FT /label=synthetic\_mutation  
FT /note="old seq (Pro)"  
XX EP370205-A.  
XX 30-MAY-1990.  
XX 28-SEP-1989; 89EP-0117981.  
XX 29-SEP-1988; 88JP-0245705.  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
XX Sasaki K, Nishi T, Yasumura S, Sato M, Itoh S;  
XX WPI; 1990-165029/22.  
XX N-PSDB; AAQ04466.  
XX Polypeptide(s) with added carbohydrate chains - formed by  
XX modification of amino acid sequence, used to improve  
XX physicochemical properties and/or activities.  
XX Disclosure; Page ?; 30pp; English.  
XX The polypeptide is a deriv. of mature urokinase, designated  
XX UK-S3 which has 2 amino acid substns. which result in an N-linked  
XX glycosylation site giving the new protein improved stability and  
XX activity.  
XX See also AAR05113-17.  
XX (Updated on 25-MAR-2003 to correct PR field.)  
XX (Updated on 25-MAR-2003 to correct PA field.)  
XX (Updated on 25-MAR-2003 to correct PI field.)  
XX Sequence 411 AA;

Query Match 100.0%; Score 793; DB 11; Length 411;

```

Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
Db 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135

RESULT 9
AAR06244
ID AAR06244 standard; protein; 411 AA.
XX
AC AAR06244;
XX
DT 07-DEC-1990 (first entry)
XX
DE Urokinase precursor protein.
XX
KW Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;
KW myocardial infarction;
XX
OS Homo sapiens.
XX
PN EP380334-A.
XX
PD 01-AUG-1990.
XX
PF 25-JAN-1990; 90EP-0300772.
XX
PR 17-MAY-1989; 89JP-0121405.
PR 27-JAN-1989; 89JP-0016406.
XX
PA (GREC ) GREEN CROSS CORP.
XX
PI Matsuda H, Ueda Y, Tamanouchi K;
XX
DR WPI; 1990-233117/31.
XX
PT Urokinase precursor-lipid composite - used as thrombolytic agent,
PT having prolonged half-life in the blood, enhanced
PT bioavailability and improved activity
XX
PS Claim 3; Fig 1; 1lpp; English.
XX
CC By forming a precursor-lipid composite, the half-life of this
CC thrombolytic agent in the blood may be increased, exhibiting
CC improved activity without abnormal acceleration of fibrinolytic
CC activity. Compound is useful as a thrombolytic agent in
CC treatment of cerebral thrombosis, myocardial infarction etc.
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 793; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
Db 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135

Query Match 100.0%; Score 793; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
Db 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135

```

```

Db 121 PLVQECMVHDCADGK 135

RESULT 10
AAR10057
ID AAR10057 standard; Protein; 411 AA.
XX
AC AAR10057;
XX
DT 18-MAR-1991 (first entry)
XX
DE Pro-urokinase derivative UK-T4 with Ser(153) and Thr(155).
XX
KW pro-urokinase; UK-T4; plasminogen activator; myocardial infarction;
KW cerebral thrombosis.
XX
OS Homo sapiens.
XX
PN EP405285-A.
XX
PD 02-JAN-1991.
XX
PF 18-JUN-1990; 90EP-0111471.
XX
PR 19-JUN-1989; 89JP-0156302.
XX
PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI Yasamura S, Nishi T, Ito S;
XX
DR WPI; 1991-008678/02.
DR N-PSDB; AAQ10169.
XX
PT New plasminogen activator almost identical to natural
PT pro-urokinase - is thrombin resistant and used for
PT prophylaxis, treatment of cerebral thrombosis or myocardial
XX infarction
PS Disclosure; Page 8; 84pp; English.
XX
CC UK-T4 is one example of a plasminogen activator which differs from
CC natural human pro-urokinase at positions 153 and 155 (Leu
CC substituted by Ser; Pro substituted by Thr, respectively).
CC The derivative has decreased susceptibility to thrombin compared to
CC natural type pro-UK and higher specific activity.
CC See also AAQ10168 and AAQ10170.
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 793; DB 12; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
Db 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135

RESULT 11
AAR10058
ID AAR10058 standard; Protein; 411 AA.
XX
AC AAR10058;
XX

```

DT 18-MAR-1991 (first entry)  
 XX Pro-proteinase derivative UK-S3 with Asn(153) and Thr(155).  
 XX pro-urokinase; UK-S3; plasminogen activator; myocardial infarction;  
 KW cerebral thrombosis.  
 XX Homo sapiens.  
 XX EP405285-A.  
 XX 02-JAN-1991.  
 PD 18-JUN-1990; 90EP-0111471.  
 XX 19-JUN-1989; 89JP-0156302.  
 XX (KYOW ) KYOWA HAKKO KOGYO KK.  
 PA Yasamura S, Nishi T, Ito S;  
 XX WPI; 1991-008678/02.  
 DR N-PSDB; AAQ10170.  
 XX New plasminogen activator almost identical to natural  
 PT pro-urokinase - is thrombin resistant and used for  
 PT prophylaxis, treatment of cerebral thrombosis or myocardial  
 PT infarction  
 XX Disclosure; Page 9; 84pp; English.  
 XX UK-S3 is one example of a plasminogen activator which differs from  
 CC natural human pro-urokinase at positions 153 and 155. (Leu  
 CC substituted by Asn; Pro substituted by Thr, respectively)  
 CC The derivative has decreased susceptibility to thrombin compared to  
 CC natural type pro-UK and higher specific activity.  
 CC See also AAQ10168 and AAQ10169.  
 XX Sequence 411 AA;  
 SQ Query Match 100.0%; Score 793; DB 12; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-51;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60  
 Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60  
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120  
 Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120  
 QY 121 PLVQECMVHDCADGK 135  
 Db 121 PLVQECMVHDCADGK 135  
 RESULT 12  
 AAR40225  
 ID AAR40225 standard; Protein; 411 AA.  
 AC AAR40225;  
 XX 10-FEB-1994 (first entry)  
 DT PUK.  
 DE Pre-urokinase; thrombolytic; blood; plasmid; PUK.  
 KW Homo sapiens.  
 XX Key Location/Qualifiers  
 XX Region 10...60  
 FT

FT /note= "Ser, Asn, Pro, Gly, or Tyr in the region  
 FT 10 to 60 in the N-terminal of the human PUK  
 FT can be replaced by Thr, Pro or Ala"  
 FT Misc-difference 64  
 FT /note= "Thr encoded by AGT (sic)"  
 FT Misc-difference 177  
 FT /note= "Thr encoded by TAC (sic)"  
 FT Domain 10..49  
 FT /label= EGF  
 FT Region 10..19  
 FT /label= Loop\_1  
 FT Region 20..31  
 FT /label= Loop\_2  
 FT Region 33..42  
 FT /label= Loop\_3  
 XX JP05192142-A.  
 PN 03-AUG-1993.  
 XX 20-JAN-1992; 92JP-0030178.  
 XX 20-JAN-1992; 92JP-0030178.  
 PR (GREC ) GREEN CROSS CORP.  
 XX WPI; 1993-277461/35.  
 DR N-PSDB; AAQ48228.  
 XX Mutant human pre-urokinase - by replacing specified aminoacid(s)  
 PT in N-terminal for providing longer half-life in blood and higher  
 PT thrombolytic ability  
 XX Claim 1; Page 14-16; 26pp; Japanese.  
 CC Human pre-urokinase can be modified by replacing: Ser, Asn,  
 CC Pro, Gly, or Tyr in the region ranging from residue 10 to 60 in the  
 CC N-terminal of the human PUK by Thr, Pro or Ala. These mutants have a  
 CC longer half-life in the blood and a higher thrombolytic ability.  
 CC For examples see (AAR47956-R47960).  
 XX Sequence 411 AA;  
 SQ Query Match 100.0%; Score 793; DB 14; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-51;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60  
 Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60  
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120  
 Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120  
 QY 121 PLVQECMVHDCADGK 135  
 Db 121 PLVQECMVHDCADGK 135  
 RESULT 13  
 AAR62991  
 ID AAR62991 standard; protein; 411 AA.  
 XX AAR62991;  
 XX 25-MAR-2003 (updated)  
 DT 21-SEP-1995 (first entry)  
 DE Pro-urokinase.  
 XX Pro-urokinase; thrombolysis; fibrin clot lysis.  
 KW  
 XX

```

OS Homo sapiens.
XX Key Location/Qualifiers
FH Domain 297..313
FT Disulfide-bond /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX 12-JAN-1995.
XX 28-JUN-1994; 94WO-US07278.
XX 02-JUL-1993; 93US-0087163.
XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX Gurewich V, Liu J;
XX WPI; 1995-060991/08.
XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation
XX Disclosure; Fig 1; 46pp; English.
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
XX described in AAR62992-R63008 were derived. These mutants retain the
XX thrombolytic activity of the wild type protein, useful for the
XX treatment of thromboembolism, but have a reduced fibrinogenolysis
XX activity and non-specific plasminogen activation. The mutants can
XX therefore be used for the lysis of fibrin clots without inducing
XX systemic bleeding, as can be the case with the wild type protein.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 411 AA;
XX Query Match 100.0%; Score 793; DB 16; Length 411;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-51;
XX Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRPNPDRRRPWCYVQVGLK 120
Db 61 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRPNPDRRRPWCYVQVGLK 120
QY 121 PLVQECWVHDCADGK 135
Db 121 PLVQECWVHDCADGK 135
RESULT 14
AAR62992
ID AAR62992 standard; protein; 411 AA.
XX AAR62992;
XX AC
XX XX
DT 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)

```

```

XX DE Pro-urokinase mutant Ala300.
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala300;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding.
XX OS Homo sapiens.
XX Key Location/Qualifiers
FH Domain 297..313
FT Disulfide-bond /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX 12-JAN-1995.
XX 28-JUN-1994; 94WO-US07278.
XX 02-JUL-1993; 93US-0087163.
XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX Gurewich V, Liu J;
XX WPI; 1995-060991/08.
XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation
XX Claim 5; Fig 1; 46pp; English.
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
XX described in AAR62992-R63008 were derived. These mutants retain the
XX thrombolytic activity of the wild type protein, useful for the
XX treatment of thromboembolism, but have a reduced fibrinogenolysis
XX activity and non-specific plasminogen activation. The mutants can
XX therefore be used for the lysis of fibrin clots without inducing
XX systemic bleeding, as can be the case with the wild type protein.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 411 AA;
XX Query Match 100.0%; Score 793; DB 16; Length 411;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-51;
XX Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRPNPDRRRPWCYVQVGLK 120
Db 61 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRPNPDRRRPWCYVQVGLK 120
QY 121 PLVQECWVHDCADGK 135
Db 121 PLVQECWVHDCADGK 135
RESULT 15

```

AAR62993		121 PLVQECMVHDCADGK 135	
ID	AAR62993 standard; protein; 411 AA.	Db	121 PLVQECMVHDCADGK 135
XX			
AC	AAR62993;		
XX			
DT	25-MAR-2003 (updated)		
DT	21-SEP-1995 (first entry)		
XX			
DE	Pro-urokinase mutant His300.		
XX			
XX	Pro-urokinase; thrombolysis; fibrin clot lysis; mutant His300;		
KW	reduced fibrinogenolysis; non-specific plasminogen activation;		
KW	systemic bleeding.		
KW			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Domain	297..313	
FT	/note= "flexible loop"		
FT	Disulfide-bond	11..19	
FT	Disulfide-bond	13..31	
FT	Disulfide-bond	33..42	
FT	Disulfide-bond	50..131	
FT	Disulfide-bond	71..113	
FT	Disulfide-bond	102..126	
FT	Disulfide-bond	148..279	
FT	Disulfide-bond	189..205	
FT	Disulfide-bond	197..268	
FT	Disulfide-bond	293..362	
FT	Disulfide-bond	325..341	
FT	Disulfide-bond	352..380	
XX			
PN	WO9501427-A1.		
XX			
PD	12-JAN-1995.		
XX			
XX	28-JUN-1994; 94WO-US07278.		
XX			
XX	02-JUL-1993; 93US-0087163.		
XX			
PA	(NEWE-) NEW ENGLAND DEACONESS HOSPITAL.		
XX			
PI	Gurewich V, Liu J;		
XX			
DR	WPI; 1995-060991/08.		
XX			
PT	Pro-urokinase mutants - have thrombolytic activity but reduced		
PT	fibrinogenolysis activity and non-specific plasminogen activation		
XX			
PS	Claim 5; Fig 1; 46pp; English.		
XX			
CC	AAR62991 is the wild type pro-urokinase, from which the new mutants		
CC	described in AAR62992-R6308 were derived. These mutants retain the		
CC	thrombolytic activity of the wild type protein, useful for the		
CC	treatment of thromboembolism, but have a reduced fibrinogenolysis		
CC	activity and non-specific plasminogen activation. The mutants can		
CC	therefore be used for the lysis of fibrin clots without inducing		
CC	systemic bleeding, as can be the case with the wild type protein.		
CC	(Updated on 25-MAR-2003 to correct PN field.)		
XX			
SQ	Sequence 411 AA;		
Query Match		100.0%; Score 793; DB 16; Length 411;	
Best Local Similarity		100.0%; Pred. No. 1.9e-51;	
Matches 135; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 SNELHQVPSNCDLNGGTCVSNKYFSNTHWCNPKKFGGQHCETDKSKTCYEGNGHFYRG 60		
DB	1 SNELHQVPSNCDLNGGTCVSNKYFSNTHWCNPKKFGGQHCETDKSKTCYEGNGHFYRG 60		
QY	61 KASTDTMGRPCLPWNASVLOQTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQGLK 120		
DB	61 KASTDTMGRPCLPWNASVLOQTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQGLK 120		



QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
 |||||||  
 Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
 |||||||  
 QY 61 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
 |||||||  
 Db 61 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
 |||||||  
 QY 121 PLVQECMVHDCADGK 135  
 |||||||  
 Db 121 PLVQECMVHDCADGK 135  
 |||||||

## RESULT 17

AAR62995  
 ID AAR62995 standard; protein; 411 AA.

XX AC AAR62995;

XX 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

XX Pro-urokinase mutant Ala300 Ala301.

DE Pro-urokinase; thrombolysis; fibrin clot lysis;  
 XX reduced fibrinogenolysis; non-specific plasminogen activation;  
 KW mutant Ala300 Ala301; systemic bleeding.  
 KW  
 KW  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Domain 297..313  
 FT /note= "flexible loop"

FT Disulfide-bond 11..19  
 FT Disulfide-bond 13..31  
 FT Disulfide-bond 33..42  
 FT Disulfide-bond 50..131  
 FT Disulfide-bond 71..113  
 FT Disulfide-bond 102..126  
 FT Disulfide-bond 148..279  
 FT Disulfide-bond 189..205  
 FT Disulfide-bond 197..268  
 FT Disulfide-bond 293..362  
 FT Disulfide-bond 325..341  
 FT Disulfide-bond 352..380

XX WO9501427-A1.

XX 12-JAN-1995.

XX 28-JUN-1994; 94WO-US07278.

XX 02-JUL-1993; 93US-0087163.

XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.

XX Gurewich V, Liu J;

XX WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced  
 PT fibrinogenolysis activity and non-specific plasminogen activation

XX Claim 9; Fig 1; 46pp; English.

XX AAR62991 is the wild type pro-urokinase, from which the new mutants  
 CC described in AAR62992-R63008 were derived. These mutants retain the  
 CC thrombolytic activity of the wild type protein, useful for the  
 CC treatment of thromboembolism, but have a reduced fibrinogenolysis  
 CC activity and non-specific plasminogen activation. The mutants can  
 CC therefore be used for the lysis of fibrin clots without inducing  
 CC systemic bleeding, as can be the case with the wild type protein.  
 CC (Updated on 25-MAR-2003 to correct FN field.)

XX SQ Sequence 411 AA;  
 Query Match 100.0%; Score 793; DB 16; Length 411;  
 Best Local Similarity 100.0%; Pred. No 1.9e-51;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
 |||||||  
 Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
 |||||||  
 QY 61 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
 |||||||  
 Db 61 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
 |||||||  
 QY 121 PLVQECMVHDCADGK 135  
 |||||||  
 Db 121 PLVQECMVHDCADGK 135  
 |||||||

## RESULT 18

AAR62996

ID AAR62996 standard; protein; 411 AA.

XX AC AAR62996;

XX 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

XX Pro-urokinase mutant His300 Ala301.

XX Pro-urokinase; thrombolysis; fibrin clot lysis;  
 KW reduced fibrinogenolysis; non-specific plasminogen activation;  
 KW mutant His300 Ala301; systemic bleeding.  
 XX  
 XX Homo sapiens.

XX Key Location/Qualifiers  
 FH Domain 297..313  
 FT /note= "flexible loop"

FT Disulfide-bond 11..19  
 FT Disulfide-bond 13..31  
 FT Disulfide-bond 33..42  
 FT Disulfide-bond 50..131  
 FT Disulfide-bond 71..113  
 FT Disulfide-bond 102..126  
 FT Disulfide-bond 148..279  
 FT Disulfide-bond 189..205  
 FT Disulfide-bond 197..268  
 FT Disulfide-bond 293..362  
 FT Disulfide-bond 325..341  
 FT Disulfide-bond 352..380

XX WO9501427-A1.

XX 12-JAN-1995.

XX 28-JUN-1994; 94WO-US07278.

XX 02-JUL-1993; 93US-0087163.

XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.

XX Gurewich V, Liu J;

XX WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced  
 PT fibrinogenolysis activity and non-specific plasminogen activation  
 XX  
 PS Claim 9; Fig 1; 46pp; English.

XX AAR62991 is the wild type pro-urokinase, from which the new mutants

CC described in AAR62992-R63008 were derived. These mutants retain the  
CC thrombolytic activity of the wild type protein, useful for the  
CC treatment of thromboembolism, but have a reduced fibrinogenolysis  
CC activity and non-specific plasminogen activation. The mutants can  
CC therefore be used for the lysis of fibrin clots without inducing  
CC systemic bleeding, as can be the case with the wild type protein.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 411 AA;  
  
Query Match 100.0%; Score 793; DB 16; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.9e-51;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60  
  
QY 61 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYOVGLK 120  
DB 61 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYOVGLK 120  
  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135  
  
RESULT 19  
AAR62997  
ID AAR62997 standard; protein; 411 AA.  
XX  
AC AAR62997;  
XX  
DT 25-MAR-2003 (updated)  
DT 21-SEP-1995 (first entry)  
XX  
DE Pro-urokinase mutant Gly306.  
XX  
KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Gly306;  
KW reduced fibrinogenolysis; non-specific plasminogen activation;  
KW systemic bleeding.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 297..313 /note= "flexible loop"  
FT Disulfide-bond 11..19  
FT Disulfide-bond 13..31  
FT Disulfide-bond 33..42  
FT Disulfide-bond 50..131  
FT Disulfide-bond 71..113  
FT Disulfide-bond 102..126  
FT Disulfide-bond 148..279  
FT Disulfide-bond 189..205  
FT Disulfide-bond 197..268  
FT Disulfide-bond 293..362  
FT Disulfide-bond 325..341  
FT Disulfide-bond 352..380  
XX  
PN WO9501427-A1.  
XX  
PD 12-JAN-1995.  
XX  
PF 28-JUN-1994; 94WO-US07278.  
XX  
PR 02-JUL-1993; 93US-0087163.  
XX  
PA (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.  
XX  
PI Gurewich V, Liu J;  
XX  
DR WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced  
PT fibrinogenolysis activity and non-specific plasminogen activation  
XX  
PS Claim 13; Fig 1; 46pp; English.  
XX  
CC AAR62991 is the wild type pro-urokinase, from which the new mutants  
CC described in AAR62992-R63008 were derived. These mutants retain the  
CC thrombolytic activity of the wild type protein, useful for the  
CC treatment of thromboembolism, but have a reduced fibrinogenolysis  
CC activity and non-specific plasminogen activation. The mutants can  
CC therefore be used for the lysis of fibrin clots without inducing  
CC systemic bleeding, as can be the case with the wild type protein.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 411 AA;  
  
Query Match 100.0%; Score 793; DB 16; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.9e-51;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60  
  
QY 61 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYOVGLK 120  
DB 61 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYOVGLK 120  
  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135  
  
RESULT 20  
AAR62998  
ID AAR62998 standard; protein; 411 AA.  
XX  
AC AAR62998;  
XX  
DT 25-MAR-2003 (updated)  
DT 21-SEP-1995 (first entry)  
XX  
DE Pro-urokinase mutant Ala313.  
XX  
KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala313;  
KW reduced fibrinogenolysis; non-specific plasminogen activation;  
KW systemic bleeding.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 297..313 /note= "flexible loop"  
FT Disulfide-bond 11..19  
FT Disulfide-bond 13..31  
FT Disulfide-bond 33..42  
FT Disulfide-bond 50..131  
FT Disulfide-bond 71..113  
FT Disulfide-bond 102..126  
FT Disulfide-bond 148..279  
FT Disulfide-bond 189..205  
FT Disulfide-bond 197..268  
FT Disulfide-bond 293..362  
FT Disulfide-bond 325..341  
FT Disulfide-bond 352..380  
XX  
PN WO9501427-A1.  
XX  
PD 12-JAN-1995.  
XX  
PF 28-JUN-1994; 94WO-US07278.  
XX

```

PR 02-JUL-1993; 93US-0087163.
XX
XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
XX
XX Gurewich V, Liu J;
XX
XX WPI; 1995-060991/08.
XX
XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation
XX
XX Claim 11; Fig 1; 46pp; English.
XX
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
XX described in AAR62992-R63008 were derived. These mutants retain the
XX thrombolytic activity of the wild type protein, useful for the
XX treatment of thromboembolism, but have a reduced fibrinogenolysis
XX activity and non-specific plasminogen activation. The mutants can
XX therefore be used for the lysis of fibrin clots without inducing
XX systemic bleeding, as can be the case with the wild type protein.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Query Match 100.0%; Score 793; DB 16; Length 411;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-51;
XX Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWNASATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVOVGLK 120
DB 61 KASTDTMGPRCLPWNASATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVOVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135
RESULT 21
AAR62999
ID AAR62999 standard; protein; 411 AA.
XX
XX AAR62999;
XX
XX 25-MAR-2003 (updated)
XX 21-SEP-1995 (first entry)
XX
XX Pro-urokinase mutant His313.
XX
XX Pro-urokinase; thrombolysis; fibrin clot lysis; mutant His313;
XX reduced fibrinogenolysis; non-specific plasminogen activation;
XX systemic bleeding.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 297..313
XX Disulfide-bond 11..19
XX Disulfide-bond 13..31
XX Disulfide-bond 33..42
XX Disulfide-bond 50..131
XX Disulfide-bond 71..113
XX Disulfide-bond 102..126
XX Disulfide-bond 148..279
XX Disulfide-bond 189..205
XX Disulfide-bond 197..268
XX Disulfide-bond 293..362
XX Disulfide-bond 325..341
XX Disulfide-bond 352..380

```

```

XX WO9501427-A1.
XX
XX 12-JAN-1995.
XX
XX 28-JUN-1994; 94WO-US07278.
XX
XX 02-JUL-1993; 93US-0087163.
XX
XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
XX
XX Gurewich V, Liu J;
XX
XX WPI; 1995-060991/08.
XX
XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation
XX
XX Claim 11; Fig 1; 46pp; English.
XX
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
XX described in AAR62992-R63008 were derived. These mutants retain the
XX thrombolytic activity of the wild type protein, useful for the
XX treatment of thromboembolism, but have a reduced fibrinogenolysis
XX activity and non-specific plasminogen activation. The mutants can
XX therefore be used for the lysis of fibrin clots without inducing
XX systemic bleeding, as can be the case with the wild type protein.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Query Match 100.0%; Score 793; DB 16; Length 411;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-51;
XX Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWNASATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVOVGLK 120
DB 61 KASTDTMGPRCLPWNASATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVOVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135
RESULT 22
AAR63000
ID AAR63000 standard; protein; 411 AA.
XX
XX AAR63000;
XX
XX 25-MAR-2003 (updated)
XX 21-SEP-1995 (first entry)
XX
XX Pro-urokinase mutant Ser175 His187.
XX
XX Pro-urokinase; thrombolysis; fibrin clot lysis;
XX reduced fibrinogenolysis; non-specific plasminogen activation;
XX systemic bleeding; mutant Ser175 His187.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 297..313
XX Disulfide-bond 11..19
XX Disulfide-bond 13..31
XX Disulfide-bond 33..42
XX Disulfide-bond 50..131
XX Disulfide-bond 71..113

```



W Pro-urokinase; thrombolysis; fibrin clot lysis;  
W reduced fibrinogenolysis; non-specific plasminogen activation;  
W systemic bleeding; mutant Ser175 His187 Gly306.

X Homo sapiens.

X Key Location/Qualifiers  
X Domain 297..313

X Disulfide-bond 11..19

X Disulfide-bond 13..31

X Disulfide-bond 33..42

X Disulfide-bond 50..131

X Disulfide-bond 71..113

X Disulfide-bond 102..126

X Disulfide-bond 148..279

X Disulfide-bond 189..205

X Disulfide-bond 197..268

X Disulfide-bond 293..362

X Disulfide-bond 325..341

X Disulfide-bond 352..380

X WO9501427-A1.

X 12-JAN-1995.

X 28-JUN-1994; 94WO-US07278.

X 02-JUL-1993; 93US-0087163.

X (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

X Gurewich V, Liu J;

X WPI; 1995-060991/08.

X Pro-urokinase mutants - have thrombolytic activity but reduced  
X fibrinogenolysis activity and non-specific plasminogen activation  
X Claim 15; Fig 1; 46pp; English.

X AAR62991 is the wild type pro-urokinase, from which the new mutants  
X described in AAR62992-R63008 were derived. These mutants retain the  
X thrombolytic activity of the wild type protein, useful for the  
X treatment of thromboembolism, but have a reduced fibrinogenolysis  
X activity and non-specific plasminogen activation. The mutants can  
X therefore be used for the lysis of fibrin clots without inducing  
X systemic bleeding, as can be the case with the wild type protein.  
X (Updated on 25-MAR-2003 to correct PN field.)

X Sequence 411 AA;

X Query Match 100.0%; Score 793; DB 16; Length 411;

X Best Local Similarity 100.0%; Pred. No. 1.9e-51;

X Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

X 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60

X 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60

X 61 KASTDTMGRPCLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120

X 61 KASTDTMGRPCLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120

X 121 PLVQECMVHDCADGK 135

X 121 PLVQECMVHDCADGK 135

X 63003 standard; protein; 411 AA.

25

3

AC AAR63003;

XX

XX 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

XX

XX Pro-urokinase mutant Ser175 His187 His313.

DE

XX Pro-urokinase; thrombolysis; fibrin clot lysis;  
KW reduced fibrinogenolysis; non-specific plasminogen activation;  
KW systemic bleeding; mutant Ser175 His187 His313.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

XX Domain 297..313

XX Disulfide-bond 11..19

XX Disulfide-bond 13..31

XX Disulfide-bond 33..42

XX Disulfide-bond 50..131

XX Disulfide-bond 71..113

XX Disulfide-bond 102..126

XX Disulfide-bond 148..279

XX Disulfide-bond 189..205

XX Disulfide-bond 197..268

XX Disulfide-bond 293..362

XX Disulfide-bond 325..341

XX Disulfide-bond 352..380

XX WO9501427-A1.

XX

XX 12-JAN-1995.

XX

XX 28-JUN-1994; 94WO-US07278.

XX

XX 02-JUL-1993; 93US-0087163.

XX

XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

XX

XX Gurewich V, Liu J;

XX

XX WPI; 1995-060991/08.

XX

XX Pro-urokinase mutants - have thrombolytic activity but reduced  
XX fibrinogenolysis activity and non-specific plasminogen activation  
XX Claim 15; Fig 1; 46pp; English.

XX

XX AAR62991 is the wild type pro-urokinase, from which the new mutants  
XX described in AAR62992-R63008 were derived. These mutants retain the  
XX thrombolytic activity of the wild type protein, useful for the  
XX treatment of thromboembolism, but have a reduced fibrinogenolysis  
XX activity and non-specific plasminogen activation. The mutants can  
XX therefore be used for the lysis of fibrin clots without inducing  
XX systemic bleeding, as can be the case with the wild type protein.  
XX (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 411 AA;

XX Query Match 100.0%; Score 793; DB 16; Length 411;

XX Best Local Similarity 100.0%; Pred. No. 1.9e-51;

XX Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60

Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120

Db 61 KASTDTMGRPCLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120

QY 121 PLVQECMVHDCADGK 135

QY 121 PLVQECMVHDCADGK 135

Db 121 PLVQECMVHDCADGK 135

RESULT 26

AAR63004

ID AAR63004 standard; protein; 411 AA.

XX AC AAR63004;

XX 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

XX DE Pro-urokinase mutant Ser175 His187 Gly306 Ala313.

XX KW Pro-urokinase; thrombolysis; fibrin clot lysis;

KW reduced fibrinogenolysis; non-specific plasminogen activation;

KW systemic bleeding; mutant Ser175 His187 Gly306 Ala313.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Domain 297..313

FT Disulfide-bond 11..19 /note= "flexible loop"

FT Disulfide-bond 13..31

FT Disulfide-bond 33..42

FT Disulfide-bond 50..131

FT Disulfide-bond 71..113

FT Disulfide-bond 102..126

FT Disulfide-bond 148..279

FT Disulfide-bond 189..205

FT Disulfide-bond 197..268

FT Disulfide-bond 293..362

FT Disulfide-bond 325..341

FT Disulfide-bond 352..380

PN WO9501427-A1.

XX 12-JAN-1995.

XX 28-JUN-1994; 94WO-US07278.

XX 02-JUL-1993; 93US-0087163.

XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

XX Gurewich V, Liu J;

XX WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced

FT fibrinogenolysis activity and non-specific plasminogen activation

XX Claim 15; Fig 1; 46pp; English.

XX AAR62991 is the wild type pro-urokinase, from which the new mutants

CC described in AAR62992-R63008 were derived. These mutants retain the

CC thrombolytic activity of the wild type protein, useful for the

CC treatment of thromboembolism, but have a reduced fibrinogenolysis

CC activity and non-specific plasminogen activation. The mutants can

CC therefore be used for the lysis of fibrin clots without inducing

CC systemic bleeding, as can be the case with the wild type protein.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 411 AA;

Query Match 100.0%; Score 793; DB 16; Length 411;

Best Local Similarity 100.0%; Pred. No. 1.9e-51;

Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60

Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCYEGNGHFYRG 60

Qy 61 KASTDTMGRCPLPWSATVLQOTYHAHRSALQLGLGKHNCRNPDNRPRPWCYVOVGLK 120

Db 61 KASTDTMGRCPLPWSATVLQOTYHAHRSALQLGLGKHNCRNPDNRPRPWCYVOVGLK 120

Qy 121 PLVQECMVHDCADGK 135

Db 121 PLVQECMVHDCADGK 135

RESULT 27

AAR63005

ID AAR63005 standard; protein; 411 AA.

XX AC AAR63005;

XX 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

XX DE Pro-urokinase mutant Ser175 His187 Ala300 Ala301 Ala313.

XX KW Pro-urokinase; thrombolysis; fibrin clot lysis;

KW reduced fibrinogenolysis; non-specific plasminogen activation;

KW systemic bleeding; mutant Ser175 His187 Ala300 Ala301 Ala313.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Domain 297..313

FT Disulfide-bond 11..19 /note= "flexible loop"

FT Disulfide-bond 13..31

FT Disulfide-bond 33..42

FT Disulfide-bond 50..131

FT Disulfide-bond 71..113

FT Disulfide-bond 102..126

FT Disulfide-bond 148..279

FT Disulfide-bond 189..205

FT Disulfide-bond 197..268

FT Disulfide-bond 293..362

FT Disulfide-bond 325..341

FT Disulfide-bond 352..380

PN WO9501427-A1.

XX 12-JAN-1995.

XX 28-JUN-1994; 94WO-US07278.

XX 02-JUL-1993; 93US-0087163.

XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

XX Gurewich V, Liu J;

XX WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced

FT fibrinogenolysis activity and non-specific plasminogen activation

XX Claim 16; Fig 1; 46pp; English.

XX AAR62991 is the wild type pro-urokinase, from which the new mutants

CC described in AAR62992-R63008 were derived. These mutants retain the

CC thrombolytic activity of the wild type protein, useful for the

CC treatment of thromboembolism, but have a reduced fibrinogenolysis

CC activity and non-specific plasminogen activation. The mutants can

CC therefore be used for the lysis of fibrin clots without inducing

CC systemic bleeding, as can be the case with the wild type protein.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 411 AA;

CC activity and non-specific plasminogen activation. The mutants can  
CC therefore be used for the lysis of fibrin clots without inducing  
CC systemic bleeding, as can be the case with the wild type protein.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 411 AA;  
Query Match 100.0%; Score 793; DB 16; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.9e-51;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEIDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHCNCRNPNRRRPMCYVOVGLK 120  
DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHCNCRNPNRRRPMCYVOVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135  
RESULT 29  
AAR63007  
ID AAR63007 standard; protein; 411 AA.  
XX  
AC AAR63007;  
XX  
DT 25-MAR-2003 (updated)  
DT 21-SEP-1995 (first entry)  
XX  
DE Pro-urokinase mutant Ser175 His187 Ala300 Ala301 His313.  
XX  
KW Pro-urokinase; thrombolysis; fibrin clot lysis;  
KW reduced fibrinogenolysis; non-specific plasminogen activation;  
KW systemic bleeding; mutant Ser175 His187 Ala300 Ala301 His313.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 297..313 /note= "flexible loop"  
FT Disulfide-bond 11..19  
FT Disulfide-bond 13..31  
FT Disulfide-bond 33..42  
FT Disulfide-bond 50..131  
FT Disulfide-bond 71..113  
FT Disulfide-bond 102..126  
FT Disulfide-bond 148..279  
FT Disulfide-bond 189..205  
FT Disulfide-bond 197..268  
FT Disulfide-bond 293..362  
FT Disulfide-bond 325..341  
FT Disulfide-bond 352..380  
PN WO9501427-A1.  
XX  
PD 12-JAN-1995.  
XX  
PF 28-JUN-1994; 94WO-US07278.  
XX  
PR 02-JUL-1993; 93US-0087163.  
XX  
PA (NEWWE-) NEW ENGLAND DEACONESS HOSPITAL.  
XX  
PI Gurewich V, Liu J;  
XX  
DR WPI; 1995-060991/08.  
XX  
PT Pro-urokinase mutants - have thrombolytic activity but reduced  
PT fibrinogenolysis activity and non-specific plasminogen activation

Query Match 100.0%; Score 793; DB 16; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.9e-51;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEIDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHCNCRNPNRRRPMCYVOVGLK 120  
DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHCNCRNPNRRRPMCYVOVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135  
RESULT 28  
AAR63006  
ID AAR63006 standard; protein; 411 AA.  
XX  
AC AAR63006;  
XX  
DT 25-MAR-2003 (updated)  
DT 21-SEP-1995 (first entry)  
XX  
DE Pro-urokinase mutant Ser175 His187 His300 Ala301 Ala313.  
XX  
KW Pro-urokinase; thrombolysis; fibrin clot lysis;  
KW reduced fibrinogenolysis; non-specific plasminogen activation;  
KW systemic bleeding; mutant Ser175 His187 His300 Ala301 Ala313.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 297..313 /note= "flexible loop"  
FT Disulfide-bond 11..19  
FT Disulfide-bond 13..31  
FT Disulfide-bond 33..42  
FT Disulfide-bond 50..131  
FT Disulfide-bond 71..113  
FT Disulfide-bond 102..126  
FT Disulfide-bond 148..279  
FT Disulfide-bond 189..205  
FT Disulfide-bond 197..268  
FT Disulfide-bond 293..362  
FT Disulfide-bond 325..341  
FT Disulfide-bond 352..380  
PN WO9501427-A1.  
XX  
PD 12-JAN-1995.  
XX  
PF 28-JUN-1994; 94WO-US07278.  
XX  
PR 02-JUL-1993; 93US-0087163.  
XX  
PA (NEWWE-) NEW ENGLAND DEACONESS HOSPITAL.  
XX  
PI Gurewich V, Liu J;  
XX  
DR WPI; 1995-060991/08.  
XX  
PT Pro-urokinase mutants - have thrombolytic activity but reduced  
PT fibrinogenolysis activity and non-specific plasminogen activation  
XX  
PS Claim 16; Fig 1; 46pp; English.  
XX  
CC AAR62991 is the wild type pro-urokinase, from which the new mutants  
CC described in AAR62992-R63008 were derived. These mutants retain the  
CC thrombolytic activity of the wild type protein, useful for the  
CC treatment of thromboembolism, but have a reduced fibrinogenolysis

Thu Dec 4 06:30:57 2003

XX Claim 16; Fig 1; 46pp; English.  
PS AAR62991 is the wild type pro-urokinase, from which the new mutants  
XX described in AAR62992-R63008 were derived. These mutants retain the  
CC thrombolytic activity of the wild type protein, useful for the  
CC treatment of thromboembolism, but have a reduced fibrinogenolysis  
CC activity and non-specific plasminogen activation. The mutants can  
CC therefore be used for the lysis of fibrin clots without inducing  
CC systemic bleeding, as can be the case with the wild type protein.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX Sequence 411 AA;  
SQ Query Match 100.0%; Score 793; DB 16; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.9e-51;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60  
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60  
Qy 61 KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
Db 61 KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
Qy 121 PLVOECMVHDCADGK 135  
Db 121 PLVOECMVHDCADGK 135  
RESULT 30  
AAR63008  
ID AAR63008 standard; protein; 411 AA.  
XX AAR63008;  
XX 25-MAR-2003 (updated)  
DT 21-SEP-1995 (first entry)  
XX Pro-urokinase mutant Ser175 His187 His300 Ala301 His313.  
XX Pro-urokinase; thrombolysis; fibrin clot lysis;  
KW reduced fibrinogenolysis; non-specific plasminogen activation;  
KW systemic bleeding; mutant Ser175 His187 His300 Ala301 His313.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Domain 297..313 /note= "flexible loop"  
FT Disulfide-bond 11..19  
FT Disulfide-bond 13..31  
FT Disulfide-bond 33..42  
FT Disulfide-bond 50..131  
FT Disulfide-bond 71..113  
FT Disulfide-bond 102..126  
FT Disulfide-bond 148..279  
FT Disulfide-bond 189..205  
FT Disulfide-bond 197..268  
FT Disulfide-bond 293..362  
FT Disulfide-bond 325..341  
FT Disulfide-bond 352..380  
XX WO9501427-A1.  
XX 12-JAN-1995.  
XX 28-JUN-1994; 94WO-US07278.  
XX 02-JUL-1993; 93US-0087163.  
XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

XX Gurewich V, Liu J;  
XX MPI; 1995-060991/08.  
XX Pro-urokinase mutants - have thrombolytic activity but reduced  
PT fibrinogenolysis activity and non-specific plasminogen activation  
XX Claim 16; Fig 1; 46pp; English.  
XX AAR62991 is the wild type pro-urokinase, from which the new mutants  
CC described in AAR62992-R63008 were derived. These mutants retain the  
CC thrombolytic activity of the wild type protein, useful for the  
CC treatment of thromboembolism, but have a reduced fibrinogenolysis  
CC activity and non-specific plasminogen activation. The mutants can  
CC therefore be used for the lysis of fibrin clots without inducing  
CC systemic bleeding, as can be the case with the wild type protein.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX Sequence 411 AA;  
SQ Query Match 100.0%; Score 793; DB 16; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.9e-51;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60  
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60  
Qy 61 KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
Db 61 KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
Qy 121 PLVOECMVHDCADGK 135  
Db 121 PLVOECMVHDCADGK 135  
Search completed: December 3, 2003, 14:39:13  
Job time : 22.2657 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:58 ; Search time 6.98276 Seconds  
(without alignments)  
1859.261 Million cell updates/sec

Title: US-09-880-503-4  
Perfect score: 793  
Sequence: 1 SNELHQVPSNCDCLNGGTGV.....QVGLKPLVQECMVHDCADGK 135

Scoring table: BLOSUP62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76.\*  
1: Piri.\*  
2: Pir2.\*  
3: Pir3.\*  
4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	793	100.0	431	1 UKHU	u-plasminogen acti
2	713.5	90.0	433	1 UKBAY	u-plasminogen acti
3	608	76.7	442	1 UKPG	u-plasminogen acti
4	579	73.0	433	1 JN0560	u-plasminogen acti
5	571	72.0	432	1 S18932	u-plasminogen acti
6	548	69.1	433	1 UPMS	u-plasminogen acti
7	335.5	42.3	477	1 A34369	t-plasminogen acti
8	335.5	42.3	477	2 JS0598	t-plasminogen acti
9	334.5	42.2	434	1 A35005	u-plasminogen acti
10	328.5	41.4	291	2 J38098	t-plasminogen acti
11	328.5	41.4	431	2 JS0599	t-plasminogen acti
12	328.5	41.4	562	1 UKHUT	t-plasminogen acti
13	321.5	40.5	559	1 A35029	t-plasminogen acti
14	315.5	39.8	559	1 A29941	t-plasminogen acti
15	310.5	39.2	477	2 JS0597	t-plasminogen acti
16	264.5	33.4	558	2 JC5878	plasma hyaluronan-
17	260.5	32.8	655	1 A46688	hepatocyte growth
18	257.5	32.5	560	1 JC4795	plasma hyaluronan-
19	257	32.4	603	2 S28941	coagulation factor
20	233	29.4	615	1 KPHU12	coagulation factor
21	223	28.1	394	2 JS0600	t-plasminogen acti
22	217	27.4	593	2 S45281	coagulation factor
23	169	21.3	685	1 A48289	neurotrophic recep
24	162.5	20.5	4548	1 S00657	apoptein(a) (BC
25	161	20.3	1420	2 A32869	apolipoprotein(a)
26	160	20.2	123	2 C61545	plasmin (EC 3.4.21
27	159	20.1	120	2 E61545	plasmin (EC 3.4.21
28	158.5	20.0	937	2 A45082	neurotrophic recep
29	157.5	19.9	460	2 B61545	plasmin (EC 3.4.21

30	156	19.7	812	1 PLBO	plasmin (EC 3.4.21
31	155	19.5	790	1 PLPG	plasmin (EC 3.4.21
32	154	19.4	89	2 A60140	plasmin (EC 3.4.21
33	154	19.4	810	1 PLHU	plasmin (EC 3.4.21
34	153.5	19.4	810	2 I46260	plasmin (EC 3.4.21
35	152.5	19.2	810	2 B30848	plasmin (EC 3.4.21
36	152	19.2	169	2 A40522	plasmin (EC 3.4.21
37	148	18.7	943	2 B45082	neurotrophic recep
38	147	18.5	812	1 PLMS	plasmin (EC 3.4.21
39	144.5	18.2	806	2 T18840	hypothetical prote
40	144	18.2	711	1 A47136	macrophage-stimula
41	141.5	17.8	455	2 A61545	plasmin (EC 3.4.21
42	141.5	17.8	728	1 JH0579	hepatocyte growth
43	141	17.8	411	2 I51285	hepatocyte growth
44	141	17.8	710	1 I51283	hepatocyte growth
45	139	17.5	716	1 A40332	macrophage-stimula

#### ALIGNMENTS

#### RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human

N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminog  
N;Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen a  
in form

C;Species: Homo sapiens (man)

C;Date: 17-Dec-1992 #sequence revision 04-Dec-1986 #text change 15-Sep-2000

C;Accession: A00931; I52209; J70102; A37561; I38102; S65783; A37562; A37563; A37564; A

R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasi, F.

Nucleic Acids Res. 13, 2759-2771, 1985

A;Title: The human urokinase-plasminogen activator gene and its promoter.

A;Reference number: A00931; MUID:85215647; PMID:2987867

A;Accession: A00931

A;Molecule type: DNA

A;Residues: 1-431 <RIC>

A;Cross-references: GB:X02419; NID:g37601; PIDN:CAA26268.1; PID:g1834524

A;Note: the authors translated the codon ATG for residue 214 as Ile

R;Nagamine, Y.; Pearson, D.; Grattan, M.

Biochem. Biophys. Res. Commun. 132, 563-569, 1985

A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcin

A;Reference number: I52209; MUID:86050639; PMID:3933505

A;Accession: I52209

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 145-161 <NAG1>

A;Cross-references: GB:K03027; NID:g340174; PIDN:AAA61257.1; PID:g340175

R;Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama

Gene 36, 183-188, 1985

A;Title: Molecular cloning of cDNA coding for human preprourokinase.

A;Reference number: J70102; MUID:86056954; PMID:2415429

A;Accession: J70102

A;Molecule type: mRNA

A;Residues: 1-213,'I',215-431 <NAG2>

A;Cross-references: GB:K03226; NID:g340155; PIDN:AA097138.1; PID:g340158; GB:D00244; N

R;Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.

Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984

A;Title: Identification and primary sequence of an unspliced human urokinase poly(A) +

A;Reference number: A37561; MUID:84272706; PMID:6589620

A;Accession: A37561

A;Molecule type: mRNA

A;Residues: 66-431 <VER>

A;Cross-references: GB:D00244; NID:g220138

R;Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Else.

DNA 4, 139-146, 1985

A;Title: Molecular cloning, sequencing, and expression in Escherichia coli of human pr.

A;Reference number: I38102; MUID:85203359; PMID:3888571

A;Accession: I38102

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-150,'W',152-213,'I',215-385,'C',387-429,'V',431 <JAC>

A;Cross-references: EMBL:X02760; NID:g35297; PIDN:CAA26535.1; PID:g35298

C;Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protease  
 F;1-20/Domain: signal sequence #status predicted <SIG>  
 F;21-431/Product: urokinase-type plasminogen activator, single chain form #status predicted <M>  
 F;21-177/Product: urokinase-type plasminogen activator chain A #status experimental <M>  
 F;31-62/Domain: EGF homology <EGF>  
 F;70-151/Domain: kringle homology <KRG>  
 F;156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <I>  
 F;179-431/Product: urokinase-type plasminogen activator chain B #status experimental <I>  
 F;179-419/Domain: trypsin homology <TRY>  
 F;31-33,33-51,53-62,70-151,91-133,122-146,168-299,209-225,217-288,313-382,345-361,372-382/Binding site: carbohydrate (Thr) (covalent) #status predicted  
 F;178-179/Cleavage site: Lys-Ile (plasmin) #status experimental  
 F;224,275,376/Active site: His, Asp, Ser #status experimental  
 F;322/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 100.0%; Score 793; DB 1; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-64;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCETDKSKTCYEGNGHFVRG 60  
 Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCETDKSKTCYEGNGHFVRG 80

Qy 61 KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
 Db 81 KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 140

Qy 121 PLVQECMVHDCADGK 135  
 Db 141 PLVQECMVHDCADGK 155

RESULT 2  
 UKBAY  
 u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon  
 C;Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)  
 C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 18-Jun-1999  
 C;Accession: S14687; S08651  
 R;AU, Y.F.T.; Wang, T.W.; Clowes, A.W.  
 Nucleic Acids Res. 18, 3411, 1990  
 A;Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator (EC 3.4.21.73); PMID:2113276  
 A;Reference number: S14687; PMID:2113276  
 A;Accession: S14687  
 A;Molecule type: mRNA  
 A;Residues: 1-433 <AUY>  
 A;Cross-references: EMBL:X51935; NID:G38130; PIDN:CAA36200.1; PID:G38131  
 C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t  
 C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine protease  
 F;1-20/Domain: signal sequence #status predicted <SIG>  
 F;21-176/Product: plasminogen activator chain A #status predicted <ACH>  
 F;30-61/Domain: EGF homology <EGF>  
 F;169-150/Domain: kringle homology <KRG>  
 F;178-433/Product: plasminogen activator chain B #status predicted <BCH>  
 F;178-421/Domain: trypsin homology <TRY>  
 F;167-298,208-224,216-287,315-384,347-363,374-402/Disulfide bonds: #status predicted  
 F;223,274,378/Active site: His, Asp, Ser #status predicted  
 F;324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.0%; Score 713.5; DB 1; Length 433;  
 Best Local Similarity 91.9%; Pred. No. 4.8e-57;  
 Matches 124; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

Qy 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCETDKSKTCYEGNGHFVRG 60  
 Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCETDKSKTCYEGNGHFVRG 79

Qy 61 KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
 Db 80 KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 139

Qy 121 PLVQECMVHDCADGK 135  
 Db 140 QRVQECMVHDCADGK 154

R;Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, K.  
 Biochim. Biophys. Acta 1293, 83-89, 1996  
 A;Title: Characterization of single chain urokinase-type plasminogen activator with a novel  
 A;Reference number: S65783; PMID:96186279; PMID:8652631  
 A;Accession: S65783  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>  
 A;Cross-references: EMBL:D11143; NID:G1311467; PIDN:BAA01919.1; PID:G1199928  
 R;Gunzler, W.A.; Zeffens, G.J.; Otting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L.  
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982  
 A;Title: The primary structure of high molecular mass urokinase from human urine.  
 A;Reference number: A37562; PMID:6754569  
 A;Accession: A37562  
 A;Molecule type: protein  
 A;Residues: 21-177 <GUN>  
 R;Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O.  
 Eur. J. Biochem. 125, 251-257, 1982  
 A;Title: Human low-molecular-weight urinary urokinase. Partial characterization and preliminary  
 A;Reference number: A37563; PMID:83003608; PMID:6749491  
 A;Accession: A37563  
 A;Molecule type: protein  
 A;Residues: 136-176;179-193, 'T', 195, 'T', 197-224 <SCH>  
 R;Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Flohe, L.  
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982  
 A;Title: The complete amino acid sequence of low molecular mass urokinase from human urine.  
 A;Reference number: A37564; PMID:83055099; PMID:6754572  
 A;Accession: A37564  
 A;Molecule type: protein  
 A;Residues: 159-410 <STE>  
 R;Kentzer, E.J.; Boko, A.; Menon, G.; Sarin, V.K.  
 Biochem. Biophys. Res. Commun. 171, 401-406, 1990  
 A;Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinant  
 A;Reference number: A35689; PMID:90365737; PMID:2393398  
 A;Accession: A35689  
 A;Molecule type: protein  
 A;Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>  
 A;Note: identification of a fucose and attempt to determine its attachment site  
 R;Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzman  
 Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990  
 A;Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell line  
 A;Reference number: A36697; PMID:91097529; PMID:2125213  
 A;Accession: A36697  
 A;Molecule type: protein  
 A;Residues: 21-34 <RAB>  
 R;Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.  
 Submitted to the Brookhaven Protein Data Bank, July 1993  
 A;Reference number: A51255; PDB:1KDU  
 A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
 R;Li, X.; Smith, R.A.G.; Dobson, C.M.  
 Biochemistry 31, 9562-9571, 1992  
 A;Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain of  
 A;Reference number: A44375; PMID:93003110; PMID:1327118  
 A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR  
 R;Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazar, A.P.; Olejniczak,  
 submitted to the Brookhaven Protein Data Bank, January 1994  
 A;Reference number: A66822; PDB:1LURK  
 A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
 R;Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;  
 submitted to the Brookhaven Protein Data Bank, July 1995  
 A;Reference number: A66058; PDB:1LWM  
 A;Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175;179-426  
 C;Comment: This enzyme is found in urine in a high molecular mass form, consisting of A  
 C;Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a  
 C;Genetics:  
 A;Gene: GDB:PLAU  
 A;Cross-references: GDB:119497; OMIM:191840  
 A;Map position: 10Q24-10Q24  
 A;Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3  
 C;Function:  
 A;Description: proteolytically activates plasminogen  
 A;Pathway: fibrinolysis  
 C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try



QY 9 SNCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFGYRGKASTDTMG 68  
 Db 29 SNCGCQNGGVCVSYKFSIRRCSPKPKFGHCEIDTSKTCYHNGGQSYRGKANTDTKG 88  
 QY 69 RPLCPWNSATVLQOYTHAHRSDALQGLGKHNYCRPNDRRPPWCYVQGLKPLVQECMV 128  
 Db 89 RPLCLAWNSPAVLQOYTHAHRSDALSGLGKHNYCRPNDRRPPWCYVQGLKPLVQECMV 148  
 QY 129 HDCADGK 135  
 Db 149 QDCSLSK 155  
 RESULT 6  
 UKMS  
 u-plasminogen activator (EC 3.4.21.73) precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 30-Sep-1997 #sequence\_revision 30-Sep-1987 #text\_change 18-Jun-1999  
 C:Accession: A29420; A24615  
 R:Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.  
 Biochemistry 26, 8270-8279, 1987  
 A:Title: The murine urokinase-type plasminogen activator gene.  
 A:Reference number: A29420; MUID:98163489; PMID:2831940  
 A:Accession: A29420  
 A:Molecule type: DNA  
 A:Residues: 1-433 <DEG>  
 A:Cross-references: GB:ML17922; NID:g202296; PIDN:AAA40539.1; PID:g202297  
 R:Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher, E.; J. Biochem. 148, 225-232, 1985  
 A:Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase  
 A:Reference number: A24615; MUID:85179474; PMID:2985383  
 A:Accession: A24615  
 A:Molecule type: mRNA  
 A:Residues: 1-433 <BEL>  
 A:Cross-references: GB:X02389; NID:g55127; PIDN:CAA26231.1; PID:g55128  
 C:Genetics:  
 A:Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3  
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
 F:32-63/Domain: EGF homology <EGF>  
 F:71-152/Domain: kringle homology <KRG>  
 F:180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
 F:180-421/Domain: trypsin homology <TRY>  
 F:169-301, 211-227, 219-290, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted  
 F:226, 277, 378/Active site: His, Asp, Ser #status predicted  
 Query Match 69.1%; Score 548; DB 1; Length 433;  
 Best Local Similarity 71.7%; Pred. No. 4.1e-42;  
 Matches 91; Conservative 13; Mismatches 23; Indels 0; Gaps 0;  
 QY 9 SNCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFGYRGKASTDTMG 68  
 Db 30 SNCGCQNGGVCVSYKFSIRRCSPKPKFGHCEIDTSKTCYHNGGQSYRGKANTDTKG 89  
 QY 69 RPLCPWNSATVLQOYTHAHRSDALQGLGKHNYCRPNDRRPPWCYVQGLKPLVQECMV 128  
 Db 90 RPLCLAWNSPAVLQOYTHAHRSDALSGLGKHNYCRPNDRRPPWCYVQGLKPLVQECMV 149  
 QY 129 HDCADGK 135  
 Db 150 HDCSLSK 156  
 RESULT 7  
 A34369  
 t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)  
 C:Species: Megaderma lyra  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A34369  
 R:Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob

J. Biol. Chem. 264, 17947-17952, 1989  
 A:Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin  
 A:Reference number: A34369; MUID:90036867; PMID:2509450  
 A:Accession: A34369  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-477 <GAR>  
 A:Cross-references: GB:J05082; NID:gl66080; PIDN:AAA31596.1; PID:gl66081  
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h  
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-36/Domain: propeptide #status predicted <PRO>  
 F:37-477/Product: plasminogen activator #status predicted <PLA>  
 F:42-79/Domain: fibronectin type I repeat homology <1FA>  
 F:87-120/Domain: EGF homology <EGF>  
 F:128-209/Domain: kringle homology <KRG>  
 F:226-471/Domain: trypsin homology <TRY>  
 F:42-72, 70-79, 87-98, 92-109, 111-120, 128-209, 149-191, 180-204, 214-345, 257-273, 265-334, 359-  
 F:272, 321, 428/Active site: His, Asp, Ser #status predicted  
 Query Match 42.3%; Score 335.5; DB 1; Length 477;  
 Best Local Similarity 46.3%; Pred. No. 6.7e-23;  
 Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;  
 QY 3 ELHQVP-----SNCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFGY 58  
 Db 78 QCHTVPVKSCSELRFCFNGGHCWQAASFSDP-VQCQPKGYTGKQCEVDTHATCYKDGQVTV 136  
 QY 59 RGKASTDTMGPRCLPWNNSATVLQOYTHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVG 118  
 Db 137 RGTWSTSEGAQCLNWNLSLTRTYNGRRSDAITLGLGHNHNYCRPNDRNNSKPWCYVVIKA 196  
 QY 119 LKPLVQECMVHDC 132  
 Db 197 SKFILEFCSPVCS 210  
 RESULT 8  
 JS0598  
 t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat  
 N:Alternate names: tissue plasminogen activator  
 C:Species: Desmodus rotundus (common vampire bat)  
 C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
 C:Accession: JS0598  
 R:Kraetschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D  
 Gene 105, 229-237, 1991  
 A:Title: The plasminogen activator family from the salivary gland of the vampire bat D  
 A:Reference number: JS0597; MUID:92039036; PMID:1937019  
 A:Accession: JS0598  
 A:Molecule type: mRNA  
 A:Residues: 1-477 <KRA>  
 A:Cross-references: GB:M63988; NID:gl66074; PIDN:AAA31593.1; PID:gl66075  
 C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h  
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-36/Domain: propeptide #status predicted <PRO>  
 F:37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>  
 F:42-79/Domain: fibronectin type I repeat homology <1FA>  
 F:87-120/Domain: EGF homology <EGF>  
 F:128-209/Domain: trypsin homology <TRY>  
 F:226-471/Domain: kringle homology <KRG>  
 F:42-72, 70-79, 87-98, 92-109, 111-120, 128-209, 149-191, 180-204, 214-345, 257-273, 265-334, 359-  
 F:185, 398/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:225-226/Cleavage site: His-Ser (plasma) #status predicted  
 F:272, 321, 428/Active site: His, Asp, Ser #status predicted  
 Query Match 42.3%; Score 335.5; DB 2; Length 477;  
 Best Local Similarity 46.3%; Pred. No. 6.7e-23;  
 Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;  
 QY 3 ELHQVP-----SNCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFGY 58  
 Db 78 QCHTVPVKSCSELRFCFNGGHCWQAASFSDP-VQCQPKGYTGKQCEVDTHATCYKDGQVTV 136



## RESULT 12

## UKHTT

t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human  
N;Alternate names: t-PA; tissue plasminogen activator  
C;Species: Homo sapiens (man)  
C;Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 08-Dec-2000  
C;Accession: A94004; A23529; J0562; A93293; S02125; A91343; A93951; A91322; A54645; I60  
R;NV, T.; Elgh, F.; Lund, B.  
Proc. Natl. Acad. Sci. U.S.A. 81, 5385-5359, 1984  
A;Title: The structure of the human tissue-type plasminogen activator gene: correlation  
A;Reference number: A94004; MUID:84298137; PMID:6089198  
A;Accession: A94004  
A;Molecule type: DNA  
A;Residues: 1-562 <NYT>  
A;Cross-references: GB:L00141  
A;Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translation  
F;Frieren Degen, S. J.; Rajput, B.; Reich, E.  
J. Biol. Chem. 261, 6972-6985, 1986  
A;Title: The human tissue plasminogen activator gene.  
A;Reference number: A23529; MUID:86196143; PMID:3009482  
A;Accession: A23529  
A;Molecule type: DNA  
A;Residues: 1-562 <DBG>  
A;Cross-references: GB:X03021; NID:G339817; PIDN:AAA98809.1; PID:G339818  
R;Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.  
Agric. Biol. Chem. 55, 1225-1232, 1991  
A;Title: Purification and characterization of tissue plasminogen activator secreted by H  
A;Reference number: J0562; MUID:91291340; PMID:1368681  
A;Accession: J0562  
A;Molecule type: mRNA  
A;Residues: 31-562 <ITA>  
A;Cross-references: DBJ:D01096; NID:Q220128; PIDN:BAA00881.1; PID:G441174  
A;Experimental source: embryonic lung fibroblast IMR-90 cells  
A;Note: part of this sequence, including the amino end of the mature protein, was confir  
R;Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett  
Nature 301, 214-221, 1983  
A;Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche  
A;Reference number: A93293; MUID:83115262; PMID:6337343  
A;Accession: A93293  
A;Molecule type: mRNA  
A;Residues: 1-562 <PEN>  
A;Cross-references: GB:L00141  
A;Experimental source: melanoma cells  
R;Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.  
Nucleic Acids Res. 16, 5695, 1988  
A;Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe  
A;Reference number: S02125; MUID:86262579; PMID:3133640  
A;Accession: S02125  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-562 <SAS>  
A;Cross-references: EMBL:X07393; NID:G37243; PIDN:CAA30302.1; PID:G37244  
A;Experimental source: fetal lung cells  
R;Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma  
FEBS Lett. 189, 145-149, 1985  
A;Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen  
A;Reference number: A91343; MUID:85285620; PMID:3896853  
A;Accession: A91343  
A;Molecule type: mRNA  
A;Residues: 1-38,'G',86-433,'E',435-562 <KAG>  
A;Experimental source: Detroit 562 cells; ATCC 138  
R;Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.  
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983  
A;Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activ  
A;Reference number: A93951; MUID:83169656; PMID:6572897  
A;Accession: A93951  
A;Molecule type: mRNA  
A;Residues: 251-358 <EDL>  
A;Experimental source: melanoma cells  
R;Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.  
Biochemistry 23, 3701-3707, 1984  
A;Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am  
differences.

A;Reference number: A90488; MUID:85000468; PMID:6433976  
A;Contents: annotation; melanoma cells, partial sequence of residues 36-562, active an  
R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.  
FEBS Lett. 169, 29-32, 1984  
A;Note: Differences between uterine and melanoma forms of tissue plasminogen activator:  
A;Reference number: A91322; MUID:84158956; PMID:6538514  
A;Accession: A91322  
A;Molecule type: protein  
A;Residues: 33-45;311-320 <POH>  
A;Experimental source: uterus  
A;Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln  
R;van Zonneveld, A. J.; Veerman, H.; Pannekoek, H.  
J. Biol. Chem. 261, 14214-14218, 1986  
A;Reference number: A37567; MUID:87033611; PMID:3021732  
A;Contents: annotation; fibrin binding site  
R;Verheijen, J. H.; Caspers, M. P. M.; Chang, G. T. G.; de Munk, G. A. W.; Pouwels, P. H.; Eng  
EMBO J. 5, 3525-3530, 1986  
A;Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen  
A;Reference number: A37568; MUID:87161761; PMID:3030730  
A;Contents: annotation; fibrin binding site  
R;Dodd, I.; Nunn, B.; Robinson, J. H.  
Thromb. Haemost. 59, 523-528, 1988  
A;Title: Isolation, identification and pharmacokinetic properties of human tissue-type  
A;Reference number: A60902; MUID:89044681; PMID:3142086  
A;Contents: annotation: novel forms of expressed recombinant t-PA  
R;Harris, T. J. R.; Patel, T.; Marston, F. A. O.; Little, S.; Emtage, J. S.; Opdenakker, G.  
Mol. Biol. Med. 3, 279-292, 1986  
A;Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its ex  
A;Reference number: A54645; MUID:86284200; PMID:3090401  
A;Accession: A54645  
A;Molecule type: mRNA  
A;Residues: 1-562 <HAR>  
A;Cross-references: GB:M15518; NID:G190031; PIDN:AAA60111.1; PID:G190032  
A;Note: parts of this sequence were confirmed by peptide sequencing  
R;Reddy, V. B.; Garramone, A. J.; Sasak, H.; Wei, C.  
DNA 6, 461-472, 1987  
A;Title: Expression of human uterine tissue-type plasminogen activator in mouse cells  
A;Reference number: 160110; MUID:88054470; PMID:2824147  
A;Accession: 160110  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-562 <RES>  
A;Cross-references: GB:M18182; NID:G340176; PIDN:AAA36800.1; PID:G340177  
R;Fisher, R.; Waller, E. K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W. D.  
J. Biol. Chem. 260, 11223-11230, 1985  
A;Title: Isolation and characterization of the human tissue-type plasminogen activator  
A;Reference number: 155232; MUID:85289338; PMID:3161893  
A;Accession: 155232  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-36 <RE2>  
A;Cross-references: GB:M11890; NID:G339837; PIDN:AAA61213.1; PID:G339839  
C;Comment: Cleavage by plasmin or trypsin produces two chains held together by a singl  
C;Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond.  
C;Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat  
C;Genetics:  
A;Gene: GDB:PLAT  
A;Cross-references: GDB:119496; OMIM:173370  
A;Map position: 8p12-8p12  
A;Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 5  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-32/Domain: propeptide #status predicted <PRO>  
F;33-562/Product: t-plasminogen activator #status experimental <MAT>  
F;33-310/Product: t-plasminogen activator chain A #status experimental <ACH>  
F;41-78/Domain: fibronectin type I repeat homology <IFI>  
F;86-119/Domain: EGF homology <EGF>  
F;127-208/Domain: kringle homology <KR1>  
F;215-296/Domain: kringle homology <KR2>  
F;311-562/Product: t-plasminogen activator chain B #status experimental <BCH>  
F;311-556/Domain: trypsin homology <TRY>  
F;41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203,215-296,236-278,267-291,299

F:152,483/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:210-311/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental  
F:310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental  
F:357,406/Active site: His, Asp #status predicted  
F:513/Active site: Ser #status experimental

Query Match 41.4%; Score 328.5; DB 1; Length 562;  
Best Local Similarity 46.3%; Pred. No. 3.3e-22;  
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQP-SNCD---CLNGCTCVSNKYFSNIHWNCPCFKGGQHCIEDKSKTCYEGNGHFY 58

DB 77 QCHSVPVKSCSEPRCFNGGTCQQALYFSDF-VCQPEGAGKCCIEDTRATCYEDGIGSY 135

QY 59 RGKASTDTMGRCPLPWNSTVLQOTYHAHRSDALQLGLGKHNYCRNPDRRPPWCYVQVG 118

DB 136 RGTWSTAESGAECINWSSALAQKPYSGRRPDAILGLGNHNYCRNPDNRDSEKWCYVFKA 195

QY 119 LKPLVQECMVHDCADG 134

DB 196 GKYTSRFTCPACSEG 211

## RESULT 13

A35029

t-plasminogen activator (EC 3.4.21.68) precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: A35029; A31597

R:Feng, P.; Ohlsson, M.; Ny, T.

J. Biol. Chem. 265, 2022-2027, 1990

A:Title: The structure of the TARA-less rat tissue-type plasminogen activator gene. Spec

A:Reference number: A35029; MUID:90130448; PMID:2105315

A:Accession: A35029

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-559 <FEN>

A:Cross-references: GB:M31197; NID:G207429; PID:AAA42261.1; PID:G207431; GB:J05226

R:Ny, T.; Leonardsson, G.; Hsueh, A.J.W.

DNA 7, 671-677, 1988

A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator

A:Reference number: A31597; MUID:99170114; PMID:3148445

A:Accession: A31597

A:Molecule type: mRNA

A:Residues: 1-379, 'K', 381-559 <NYT>

A:Cross-references: GB:M23697; NID:G530159; PID:AAA41812.1; PID:G530160

C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom

C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-29/Domain: propeptide #status predicted <PRO>

F:30-559/Product: t-plasminogen activator #status predicted <MAT>

F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>

F:38-75/Domain: fibronectin type I repeat homology <1F1>

F:83-116/Domain: EGF homology <EGF>

F:124-205/Domain: kringle homology <KR1>

F:213-294/Domain: kringle homology <KR2>

F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>

F:309-553/Domain: trypsin homology <TRY>

F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4

F:149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted

F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 40.5%; Score 321.5; DB 1; Length 559;

Best Local Similarity 45.6%; Pred. No. 1.4e-21;

Matches 62; Conservative 14; Mismatches 55; Indels 5; Gaps 2;

QY 3 ELHQP-----SNCDLNGCTCVSNKYFSNIHWNCPCFKGGQHCIEDKSKTCYEGNGHFY 58

DB 74 QCHSVPVKSCSEPRCFNGGTCQQALYFSDF-VCQCPDGFVGRKCDIDTRATCFEGQGIT 132

QY 59 RGKASTDTMGRCPLPWNSTVLQOTYHAHRSDALQLGLGKHNYCRNPDRRPPWCYVQVG 118

DB 133 RGTWSTAENGAEICINWSSALSQKPYSAARRPNAIKLGLGNHNYCRNPDRDVKPCYVFKA 192

QY 119 LKPLVQECMVHDCADG 134

DB 193 GKYTTEFCSTPACPKG 208

## RESULT 14

A29941

t-plasminogen activator (EC 3.4.21.68) precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: A29941; S48205; S48207; S48206

R:Richles, R.J.; Darrow, A.L.; Strickland, S.

J. Biol. Chem. 263, 1563-1569, 1988

A:Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator r

A:Reference number: A29941; MUID:88087303; PMID:2826484

A:Accession: A29941

A:Molecule type: mRNA

A:Residues: 1-559 <RIC>

A:Cross-references: GB:J03520; NID:G202109; PID:AAA40470.1; PID:G202110

R:Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.

Eur. J. Biochem. 224, 863-871, 1994

A:Title: Characterization of the murine plasma fibrinolytic system.

A:Reference number: S48202; MUID:95010076; PMID:7523120

A:Accession: S48205

A:Molecule type: protein

A:Residues: 33-37, 'X', 39-40 <LIJ>

A:Accession: S48207

A:Molecule type: protein

A:Residues: 309-316 <LI2>

A:Accession: S48206

A:Molecule type: protein

A:Residues: 33-37, 'X', 39-40 <LIW>

C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h

C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-29/Domain: propeptide #status predicted <PRO>

F:30-559/Product: t-plasminogen activator #status predicted <MAT>

F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>

F:83-116/Domain: EGF homology <EGF>

F:124-205/Domain: kringle homology <KR1>

F:213-294/Domain: kringle homology <KR2>

F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>

F:309-553/Domain: trypsin homology <TRY>

F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297

F:149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted

F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 39.8%; Score 315.5; DB 1; Length 559;

Best Local Similarity 44.5%; Pred. No. 4.9e-21;

Matches 61; Conservative 15; Mismatches 56; Indels 5; Gaps 2;

QY 3 ELHQP-----SNCDLNGGTCVSNKYFSNIHWNCPCFKGGQHCIEDKSKTCYEGNGHFY 58

DB 74 QCHSVPVKSCSEPRCFNGGTCQQALYFSDF-VCQCPDGFVGRKCDIDTRATCFEQQITY 132

QY 59 RGKASTDTMGRCPLPWNSTVLQOTYHAHRSDALQLGLGKHNYCRNPDRRPPWCYVQVG 118

DB 133 RGTWSTAESGAECINWSSVLKPYNARRPNAIKLGLGNHNYCRNPDRDLKPCYVFKA 192

QY 119 LKPLVQECMVHDCADG 135

DB 193 GKYTTEFCSTPACPKG 209

## RESULT 15

JS0597

t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat

N:Alternate names: tissue plasminogen activator

C:Species: Desmondus rotundus (common vampire bat)



C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
 C;Accession: J50597  
 R;Kraatzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Dor  
 Gene 105, 229-237, 1991  
 A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des  
 A;Reference number: J50597; MUID:92039036; PMID:1937019  
 A;Accession: J50597  
 A;Molecule type: mRNA  
 A;Residues: 1-477 <KRA>  
 A;Cross-references: GB:M63987; NID:g166070; PIDN:AAA31591.1; PID:g166071  
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F;1-21/Domain: signal sequence #status predicted <SIG>  
 F;22-36/Domain: propeptide #status predicted <PRO>  
 F;37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>  
 F;42-79/Domain: fibronectin type I repeat homology <1FA>  
 F;87-120/Domain: EGF homology <EGF>  
 F;128-209/Domain: kringle homology <KRG>  
 F;226-471/Domain: trypsin homology <TRY>  
 F;432-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4  
 F;153,398/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;225-226/Cleavage site: His-Ser (plasmin) #status predicted  
 F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 39.2%; Score 310.5; DB 2; Length 477;  
 Best Local Similarity 45.5%; Pred. No. 1.2e-20;  
 Matches 60; Conservative 14; Mismatches 53; Indels 5; Gaps 3;

Qy 5 HQVPSN-CD---CLNGGTCVSNKYFSNIHNCNPKFGGQHCEIDKSKTCYEGNGHYRFG 60  
 Db 80 HTVPVNSCEPRFCNGGTCQAVFSDP-VQCQAGYTKRCEVDTRATCYEGGQVIYRG 138  
 Qy 61 KASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLK 120  
 Db 139 TWSTAERVEICINWSSILLTRTYNGRMPDAFNLGLGNHNYCRNPNGAPKFCWYIRAGK 198  
 Qy 121 PLVQECMVHDC 132  
 Db 199 FTSESCVPVCS 210

RESULT 16  
 Plasma hyaluronan-binding protein precursor - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 11-Mar-1998 #sequence\_revision 11-Mar-1998 #text\_change 16-Jul-1999  
 C;Accession: J50578  
 R;Hashimoto, K.; Tobe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, N.  
 Biol. Pharm. Bull. 20, 1127-1130, 1997  
 A;Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-bin  
 A;Reference number: J50578; MUID:98065239; PMID:9401717  
 A;Accession: J50578  
 A;Molecule type: mRNA  
 A;Residues: 1-558 <HAS>  
 C;Comment: This protein acts as serine protease.  
 C;Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsi  
 F;1-23/Domain: signal sequence #status predicted <STG>  
 F;24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MATL>  
 F;75-106/Domain: EGF homology <EG1>  
 F;113-145/Domain: EGF homology <EG2>  
 F;152-185/Domain: EGF homology <EG3>  
 F;192-274/Domain: kringle homology <KRI>  
 F;312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MATS  
 F;312-548/Domain: trypsin homology <TRY>

Query Match 33.4%; Score 264.5; DB 2; Length 558;  
 Best Local Similarity 42.4%; Pred. No. 2e-16;  
 Matches 51; Conservative 18; Mismatches 48; Indels 3; Gaps 3;

Qy 13 CLNGGTCVSNKYFSNIHNCNPKFGGQHCEIDKSKTCYEGNGHYRFGKASTDTMGRCPL 72  
 Db 157 CONGGVCSRRHRRSRF-TCACPDQYKGFCEIGPD-DCYVGGVSYRGKSKVTQNQNCL 214

Qy 73 PWSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLKPLVQE-CMVHDC 131  
 Db 215 YNNSHLLQETYNMFMEDEATHGIAEHNFPCRDGDHDKMCFVKVNSEKWKVEYCDVTVC 274

RESULT 17  
 A46688  
 hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human  
 C;Species: Homo sapiens (man)  
 C;Date: 21-Sep-1993 #sequence\_revision 25-Aug-1995 #text\_change 08-Dec-2000  
 C;Accession: A46688  
 R;Miyazawa, K.; Shinomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.  
 J. Biol. Chem. 268, 10024-10028, 1993  
 A;Title: Molecular cloning and sequence analysis of the cDNA for a human serine protea  
 d coagulation factor XII.  
 A;Reference number: A46688; MUID:93252878; PMID:7683665  
 A;Accession: A46688  
 A;Molecule type: mRNA  
 A;Residues: 1-655 <MTY>  
 A;Cross-references: DDBJ:D14012; NID:g219680; PIDN:BAR03113.1; PID:g219681  
 A;Experimental source: liver (mRNA); serum (protein)  
 A;Note: sequence extracted from NCBI backbone (NCBIN:131227, NCBIPI:131228)  
 C;Note: parts of the sequence, including the amino ends of the heavy and light chains,  
 C;Genetics:  
 A;Gene: GDB:HGFAC; HGFA; HGFAP  
 A;Cross-references: GDB:9954514  
 A;Map position: 4p16-4p16  
 C;Function:  
 A;Description: activates hepatocyte growth factor by specific proteolytic cleavage  
 A;Pathway: tissue repair and regeneration  
 C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homolog  
 C;Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase  
 F;1-34/Domain: signal sequence #status predicted <SIG>  
 F;108-148/Domain: fibronectin type II repeat homology <1F2>  
 F;164-197/Domain: EGF homology <EG1>  
 F;202-237/Domain: fibronectin type I repeat homology <1F1>  
 F;245-278/Domain: EGF homology <EG2>  
 F;286-367/Domain: kringle homology <KRG>  
 F;373-407/Product: hepatocyte growth factor activator light chain #status experimental  
 F;408-655/Product: hepatocyte growth factor activator heavy chain #status experimental  
 F;408-641/Domain: trypsin homology <TRY>  
 F;40,48,290,468,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;164-175,169-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-  
 F;447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 32.8%; Score 260.5; DB 1; Length 655;  
 Best Local Similarity 43.7%; Pred. No. 5.1e-16;  
 Matches 52; Conservative 8; Mismatches 42; Indels 17; Gaps 2;

Qy 5 HQVPSNCDCLNGGTCVSNKYFSNIHW-----CNCPKFGGQHCEIDKSKTCYEGNGH 56  
 Db 242 HTACLSFCLNGGTC-----HLIVATGTTVCAPPGPAGRLCNIEPDERFLNGT 292  
 Qy 57 FYRGKASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCY 115  
 Db 293 GYRGVASTASGLSCLAWSLLYQELHVDVSGAALLGLGPHAYCRNPDRPWCY 351

RESULT 18  
 JC4795  
 plasma hyaluronan-binding protein precursor - human  
 N;Alternate names: hepatocyte growth factor activator-like protein; PHBP  
 N;Contains: serine proteinase (EC 3.4.21.-)  
 C;Species: Homo sapiens (man)  
 C;Date: 15-Oct-1995 #sequence\_revision 16-Aug-1996 #text\_change 19-Jul-2002  
 C;Accession: JC4795  
 R;Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tomita, M.  
 J. Biochem. 119, 1157-1165, 1996  
 A;Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP  
 r activator.  
 A;Reference number: JC4795; MUID:96425001; PMID:8827452  
 A;Accession: JC4795  
 A;Molecule type: mRNA



A;Residues: 1-560 <CHO>  
A;Cross-references: GB:S83182; NID:gl836158; PIDN:AAB46909.1; PID:gl836159  
A;Experimental source: plasma  
A;Note: parts of this sequence, including the amino ends of the mature chains, were detected by mass spectrometry.  
C;Genetics:  
A;Gene: GDB:HABP2; HABP: PHBP; HGFBAL  
A;Cross-references: GDB:4573962  
C;Complex: a disulfide-bonded heterodimer of chains produced from the same precursor; the heavy chain is a hyaluronan-binding protein; EGF homology; kringlike homology; tryptophan-rich repeat domain; chondroitin sulfate proteoglycan; glycoprotein; hyaluronidase; hyaluronan-binding protein; signal sequence #status predicted <SIG>  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-313/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>  
F;77-108/Domain: EGF homology <EGF>  
F;115-147/Domain: EGF homology <EGF>  
F;154-187/Domain: EGF homology <EGF>  
F;194-276/Domain: kringlike homology <KRI>  
F;314-550/Domain: trypsin homology <TRY>  
F;314-516/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted <CAT>  
F;54-207/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;77-88,82-97,99-108,115-125,120-136,138-147,154-165,159-176,178-187,194-276,215-257,246-257,259-276,280-297,299-313,315-326,328-341,343-359,361-374,376-389,391-404,406-419,421-434,436-449,451-464,466-479,481-494,496-509/Active site: His, Asp, Ser #status predicted  
F;362,405,509/Active site: His, Asp, Ser #status predicted  
Query Match 32.5%; Score 257.5; DB 1; Length 560;  
Best Local Similarity 42.1%; Pred. No. 8.4e-16;  
Matches 51; Conservative 18; Mismatches 49; Indels 3; Gaps 3;  
QY 13 CLNGGTCVSNKYFSNIHWCNCPKFGQHCIEIDKSKTCYEGNGHFRGKASTDTMGRPCL 72  
DB 159 CQNGATCSRHRKSKF-TCACPQPKGKFCIG-SDDCVYDGDGYSYRGKNRTVQHGACL 216  
QY 73 PWSNATVLOQTYHAHRSQALQGLGHKHNCRPNDRRPPWCYQVQGLKPLVQECMVHDC 131  
DB 217 YWNSHLLQNNYWFVEDASTHGEHNFNCPDAKPCFKVTNDKVKWEYCDVSAC 276  
QY 132 A 132  
DB 277 S 277  
RESULT 19  
S28941  
coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)  
N;Alternate names: Hageman factor  
C;Species: Cavia porcellus (guinea pig)  
C;Date: 25-Feb-1994 #sequence\_revision 03-Aug-1995 #text\_change 21-Jan-2000  
C;Accession: S28941  
R;Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.; Okabe, H.; Biochim. Biophys. Acta 1159, 113-121, 1992  
A;Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage site  
A;Reference number: S28941; MUID:93003367; PMID:1390917  
A;Accession: S28941  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-603 <SEM>  
A;Cross-references: EMBL:X68615; NID:g49578; PIDN:CAA48600.1; PID:g49579  
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;  
C;Keywords: hydrolase; serine proteinase  
F;46-87/Domain: fibronectin type II repeat homology <IF2>  
F;134-169/Domain: fibronectin type I repeat homology <FI1>  
F;177-208/Domain: EGF homology <EGF>  
F;216-294/Domain: kringlike homology <KRG>  
F;359-597/Domain: trypsin homology <TRY>  
Query Match 32.4%; Score 257; DB 2; Length 603;  
Best Local Similarity 40.8%; Pred. No. 9.9e-16;  
Matches 49; Conservative 17; Mismatches 46; Indels 8; Gaps 3;  
QY 13 CLNGGTCVSNKYFSNIHWCNCPKFGQHCIEIDKSKTCYEGNGHFRGKASTDTMGRPCL 72  
DB 182 CLNGRCLE---VEGHLLCDPCWGYTGFCDLTATSCYEGRGVSGMARTTVSGAKCQ 238  
QY 73 PWSNATVLOQTYHAHRSQALQGLGHKHNCRPNDRRPPWCYQVQGLKPLVQECMVHDC 131

Db 239 RWAS-----EATYRNMTAEQALRRGLGHHTFCRPNDRPWCFFWMGNRLSWEYCDLAQC 294  
RESULT 20  
KFHUI2  
coagulation factor XIIa (EC 3.4.21.38) precursor [validated] - human  
N;Alternate names: Hageman factor (activated)  
C;Species: Homo sapiens (man)  
C;Date: 27-Nov-1985 #sequence\_revision 30-Jun-1991 #text\_change 08-Dec-2000  
C;Accession: A29411; A26814; A00930; A25191; A22248; A21037  
J;Cool, D.E.; MacGillivray, R.T.A.  
J. Biol. Chem. 262, 13662-13673, 1987  
A;Title: Characterization of the human blood coagulation factor XII gene. Intron/exon structure  
A;Reference number: A29411; MUID:88007593; PMID:2888762  
A;Accession: A29411  
A;Molecule type: DNA  
A;Residues: 1-615 <COO>  
A;Cross-references: GB:M17466; GB:J02807; NID:gl80355; PIDN:AAB59490.1; PID:gl80357  
R;Tripos, M.; Citarella, F.; Guida, S.; Galeffi, P.; Fantoni, A.; Cortese, R. Nucleic Acids Res. 14, 3146, 1986  
A;Title: cDNA sequence coding for human coagulation factor XII (Hageman).  
A;Reference number: A26814; MUID:86176794; PMID:3754331  
A;Accession: A26814  
A;Molecule type: mRNA  
A;Residues: 4-615 <TRI>  
R;Cool, D.E.; Edgell, C.J.S.; Louie, G.V.; Zoller, M.J.; Brayer, G.D.; MacGillivray, R. J. Biol. Chem. 260, 13666-13676, 1985  
A;Title: Characterization of human blood coagulation factor XII cDNA. Prediction of the structure  
A;Reference number: A00930; MUID:86033830; PMID:3877053  
A;Accession: A00930  
A;Molecule type: mRNA  
A;Residues: 14-332, 'S', 334-615 <CO2>  
A;Cross-references: GB:M11723; NID:gl80358; PIDN:AAAS1986.1; PID:gl80359  
R;Que, B.G.; Davie, E.W. Biochemistry 25, 1525-1528, 1986  
A;Title: Characterization of a cDNA coding for human factor XII (Hageman factor).  
A;Reference number: A25191; MUID:86216049; PMID:3011063  
A;Accession: A25191  
A;Molecule type: mRNA  
A;Residues: 146-378, 'G', 380-615 <QUE>  
R;McMullen, B.A.; Fujikawa, K. J. Biol. Chem. 260, 5328-5341, 1985  
A;Title: Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated factor XIIa)  
A;Reference number: A22248; MUID:85182674; PMID:3886654  
A;Accession: A22248  
A;Molecule type: protein  
A;Residues: 20-379 <WCM>  
R;Fujikawa, K.; McMullen, B.A. J. Biol. Chem. 258, 10924-10933, 1983  
A;Title: Amino acid sequence of human beta-factor XIIa.  
A;Reference number: A21037; MUID:83291041; PMID:6604055  
A;Accession: A21037  
A;Molecule type: protein  
A;Residues: 354-362, 373-615 <FUJ>  
R;Harris, R.J.; Ling, V.T.; Spellman, M.W. J. Biol. Chem. 267, 5102-5107, 1992  
A;Title: O-linked fucose is present in the first epidermal growth factor domain of factor XIIa  
A;Reference number: A44606; MUID:92184750; PMID:1544894  
A;Contents: annotation; carbohydrate binding site  
C;Genetics:  
A;Gene: GDB:FI2  
A;Cross-references: GDB:119892; OMIM:234000  
A;Map position: 5q34-5qter  
C;Introns: 19/3; 39/1; 72/2; 96/1; 133/1; 177/1; 212/1; 267/2; 340/1; 417/2; 463/1; 511/1  
C;Complex: factor XII, prekallikrein, and HWW kininogen form a complex bound to anionic site  
C;Function:  
A;Description: factor XIIa catalyzes the proteolytic activation of plasminogen, plasma kallikrein  
A;Pathway: blood coagulation; fibrinolysis  
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;  
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; kringlike; plasma;

Thu Dec 4 06:30:58 2003

```
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-372,373-615/Product: coagulation factor XIIa, alpha form #status experimental <A12>
F;47-88/Domain: fibronectin type II repeat homology <FB2>
F;98-130/Domain: EGF homology <EG1>
F;135-170/Domain: fibronectin type I repeat homology <1F1>
F;178-209/Domain: EGF homology <EG2>
F;217-295/Domain: kringle homology <KR>
F;298-356/Region: proline-rich
F;373-609/Domain: trypsin homology <TRY>
F;98-110,104-119,121-130,135-163,161-170,178-189,183-198,200-209,217-295,238-277,266-290
F;109/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;249,433/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;239,305,328,329,337/Binding site: carbohydrate (Thr) (covalent) #status predicted
F;308/Binding site: carbohydrate (Ser) (covalent) #status predicted
F;412,461,563/Active site: His, Asp, Ser #status predicted

Query Match          29.4%; Score 233; DB 1; Length 615;
Best Local Similarity 42.3%; Pred. No. 1.5e-13;
Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;

Qy 13 CLNGGTCVSNKYFSNIHWCNPKFGQHCIDKSKTCYEGNGHFGYRGKASTDTMGPRCL 72
Db 183 CLHGRCLE---VEGHLRCHCPVGVGTGPFCDVDTKASCDYGRGLSYRGLARTTLGAPCQ 239

Qy 73 PWSATVLQQTYY-HAHRSDALQGLGKHNKNCNPNDRRRPCVY 115
Db 240 PWAS-----EATYVNTAEQVNLWGLGHAFCRNPDIRPCFV 279

RESULT 21
JS0600
N-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
t-Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0600
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A;Reference number: JS0597; MUID:92039036; PMID:1937019
A;Accession: JS0600
A;Molecule type: mRNA
A;Residues: 1-394 <KRA>
A;Cross-references: GB:M63990; NID:g166078; PIDN:AAA31595.1; PID:g166079
A;Note: the authors translated the codon ATC for residue 75 as Thr
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-394/Product: plasminogen activator gamma #status predicted <PLA>
F;45-126/Domain: kringle homology <KR>
F;143-388/Domain: trypsin homology <TRY>
F;45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:
F;142-143/Cleaveage site: His-Ser (plasmin) #status predicted
F;189,238,345/Active site: His, Asp, Ser #status predicted
F;1315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          28.1%; Score 223; DB 2; Length 394;
Best Local Similarity 44.3%; Pred. No. 8.1e-13;
Matches 39; Conservative 13; Mismatches 36; Indels 0; Gaps 0;

Qy 45 DSKSKTCYEGNGHFGYRGKASTDTMGPRCLPWSATVLQQTYYHAHRSDALQGLGKHNKNCN 104
Db 40 DPHATYKDDQGVYGTGTSSTSSGACQINWNLIRTYGRMPPEAVKLGIGNHNKNCN 99

Qy 105 PNNRRRPWCYVQVGLKPLVQECWHDCA 132
Db 100 PGASKPWCYVVKARFTSESQVPCS 127

RESULT 22
S45281
```

```
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)
N;Alternate names: Hageman factor (activated)
C;Species: Bos primigenius taurus (Cattle)
C;Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C;Accession: S45281; A61329
R;Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A;Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): coi
A;Reference number: S45281; MUID:94242782; PMID:8186251
A;Accession: S45281
A;Molecule type: mRNA
A;Residues: 1-593 <SHI>
A;Cross-references: GB:S70164
A;Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70
s Pro, CTC for residue 203 as Phe, GIG for residue 247 as Leu, CCG for residue 286 as
is, and ATC for residue 505 as Leu
R;Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A;Title: Isolation and characterization of bovine factor XII (Hageman factor).
A;Reference number: A61329; MUID:77182112; PMID:861210
A;Accession: A61329
A;Molecule type: protein
A;Residues: 10-16,'X',18-19;525-550 <FUJ>
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homolog
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;
F;37-78/Domain: fibronectin type II repeat homology <1F2>
F;86-120/Domain: EGF homology <EGF>
F;125-160/Domain: fibronectin type I repeat homology <FB1>
F;207-287/Domain: kringle homology <KR>
F;350-587/Domain: trypsin homology <TRY>
F;541/Active site: Ser #status predicted

Query Match          27.4%; Score 217; DB 2; Length 593;
Best Local Similarity 35.7%; Pred. No. 4e-12;
Matches 46; Conservative 15; Mismatches 58; Indels 10; Gaps 4;

Qy 6 QVPSNCCLNGGTCVSNKYFSNIHWCNPKFGQHCIDKSKTCYE--GNGHFYRGKAS 63
Db 166 QVCTNPCLNGDSCLOAE---GHLRCPASFAGLCDVDLKASCYDDRDGLSVRGWAG 222

Qy 64 TDTMGPRCLPWSATVLQQTYY-HAHRSDALQGLGKHNKNCNPNDRRPWCYVQVGLKPL 122
Db 223 TTLGAPCQSWAS----EATYVNTAEQVNLWGLGHAFCRNPDIRPCFVIMKGDRLS 278

Qy 123 VQECWHDCA 131
Db 279 WNYCRLAPC 287

RESULT 23
A48289
neurotrophic receptor ror precursor - fruit fly (Drosophila melanogaster)
N;Alternate names: trk-related receptor
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 17-Nov-2000
C;Accession: A48289
R;Wilson, C.; Gøberdhan, D.C.I.; Steller, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 7109-7113, 1993
A;Title: Dror, a potential neurotrophic receptor gene, encodes a Drosophila homolog of
A;Reference number: A48289; MUID:93348222; PMID:8394009
A;Accession: A48289
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-685 <WIL>
A;Cross-references: GB:L20297; NID:g348103; PIDN:AAA28860.1; PID:g348104
C;Genetics:
A;Gene: FlyBase:bsk
A;Cross-references: FlyBase:FBgn0010407
C;Superfamily: Drosophila neurotrophic receptor ror; kringle homology; protein kinase
C;Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyr
F;237-310/Domain: kringle homology <KR>
F;314-338/Domain: transmembrane #status predicted <TM1>
```

```

A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-16 <R3>
A:A:Cross-references: GB:M68677; NID:gl78780; PIDN:ABA49909.1; PID:G553185
A:Note: apo(a) gene 1 (nomenclature of reference I52415)
A:Accession: I65286
A:A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-16 <R4>
A:A:Cross-references: GB:M68678; NID:gl78782; PIDN:AAA51749.1; PID:G553186
C:Genetics:
A:Gene: GDB:LPA
A:A:Cross-references: GDB:120699; OMIM:152200
A:Map position: 6q26-6q27
A:Note: several genes closely linked on chromosome 6 are identical in the first coding
rs of kringle repeats
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-4548/Product: apolipoprotein(a) #status experimental <MAT>
F:28-105/Domain: kringle homology <KR1>
F:142-219/Domain: kringle homology <KR2>
F:256-333/Domain: kringle homology <KR3>
F:370-447/Domain: kringle homology <KR4>
F:484-561/Domain: kringle homology <KR5>
F:598-675/Domain: kringle homology <KR6>
F:712-789/Domain: kringle homology <KR7>
F:826-903/Domain: kringle homology <KR8>
F:940-1017/Domain: kringle homology <KR9>
F:1054-1131/Domain: kringle homology <KR10>
F:1168-1245/Domain: kringle homology <KR11>
F:1282-1359/Domain: kringle homology <KR12>
F:1396-1473/Domain: kringle homology <KR13>
F:1510-1587/Domain: kringle homology <KR14>
F:1624-1701/Domain: kringle homology <KR15>
F:1738-1815/Domain: kringle homology <KR16>
F:1852-1929/Domain: kringle homology <KR17>
F:1966-2043/Domain: kringle homology <KR18>
F:2080-2157/Domain: kringle homology <KR19>
F:2194-2271/Domain: kringle homology <KR20>
F:2308-2385/Domain: kringle homology <KR21>
F:2422-2499/Domain: kringle homology <KR22>
F:2536-2613/Domain: kringle homology <KR23>
F:2650-2727/Domain: kringle homology <KR24>
F:2764-2841/Domain: kringle homology <KR25>
F:2878-2955/Domain: kringle homology <KR26>
F:2992-3069/Domain: kringle homology <KR27>
F:3106-3183/Domain: kringle homology <KR28>
F:3220-3297/Domain: kringle homology <KR29>
F:3334-3411/Domain: kringle homology <KR30>
F:3448-3525/Domain: kringle homology <KR31>
F:3562-3639/Domain: kringle homology <KR32>
F:3676-3753/Domain: kringle homology <KR33>
F:3782-3859/Domain: kringle homology <KR34>
F:3896-3973/Domain: kringle homology <KR35>
F:4010-4087/Domain: kringle homology <KR36>
F:4124-4201/Domain: kringle homology <KR37>
F:4228-4307/Domain: kringle homology <KR38>
F:4328-4541/Domain: trypsin homology <TRY>

Query Match 20.5%; Score 162.5; DB 1; Length 4548;
Best Local Similarity 33.1%; Pred. No. 1.8e-06;
Matches 41; Conservative 12; Mismatches 46; Indels 25; Gaps 7;

QY 27 NIHW--CN---CP-----KKFGGQHCIDKSKT---CYEGNGHFYRGKASTDTMGR 69
Db 3742 NVRWECYCNLTQCPVTESSVLATSTAVSEQAPTEQSFQDCYHGDGQSYRGSFSTVTGR 3801

QY 70 PCLPWNASATVLQOOTHARSDALQLGKHNYCRNPDNRRRPWCYVQVGLPLV--QECM 127
Db 3802 TCQSSWMT---PHWHQRTTEYPNGGLTRNYCRNPDAAIEIRPCVCT---MDPFSVRWBYCN 3855

QV 128 VHDC 131

```

[illegible][illegible]

QY 42 C-----EIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDAL 92  
Db 296 CIRIGIPMADPINKNHKCYNSTGDYRGTVSVTKSGRQCPWNS-----QYPHTHTFTAL 350  
QY 93 QLG--LGKYNICRNPDRRR--PWCY 114  
Db 351 RFPELNGGHSYCRNPGNQKEAPWC 375

RESULT 29

plasmin (EC 3.4.21.7) precursor - sheep (fragments)  
N;Alternate names: plasminogen  
N;Contains: miniplasminogen  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 28-Oct-1994 #sequence\_revision 01-Nov-1996 #text\_change 17-Mar-1999  
C;Accession: B61545; S28200  
R;Schaller, J.; Rickli, E.E.  
Enzyme 40, 63-69, 1988  
A;Title: Structural aspects of the plasminogen of various species.  
A;Reference number: A61545; MUID:89005015; PMID:3168975  
A;Accession: B61545  
A;Molecule type: protein  
A;Residues: 1-37;38-117 <SCH>  
R;Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.  
Protein Seq. Data Anal. 5, 21-25, 1992  
A;Title: Complete amino acid sequence of ovine miniplasminogen.  
A;Reference number: S28200; MUID:93149995; PMID:1492092  
A;Accession: S28200  
A;Molecule type: protein  
A;Residues: 118-460 <SC2>  
C;Superfamily: plasmin; kringlike homology; plasminogen-related protein precursor homology  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringie; plasmin; plasminogen; plasminogenase; <PRO>  
F;1-37/Domain: activation peptide (fragment) #status experimental <PRO>  
F;38-117,118-230,231-460/Product: plasmin (fragments) #status experimental <APT>  
F;41-118/Domain: kringie homology <KR4>  
F;118-460/Product: miniplasminogen #status experimental <MIN>  
F;132-211/Domain: kringie homology <KR5>  
F;226-460/Domain: plasmin chain B #status experimental <BCH>  
F;231-453/Domain: trypsin homology <TRY>  
F;272,315,410/Active site: His, Asp, Ser #status predicted

Query Match 19.9%; Score 157.5; DB 2; Length 460;  
Best Local Similarity 38.3%; Pred. No. 7.5e-07;  
Matches 34; Conservative 13; Mismatches 27; Indels 15; Gaps 4;

QY 35 KFGGQ-----HCEIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATV--LQOTYH 85  
Db 20 KKLGRSVEDCAKCE--EEAQDCYHNGQGYRGTSVTVTGRKQCQSWSSIPHRHOKTPE 78  
QY 86 AHRSDALQLGLGKHNYCRNPDRRRPWCY 114  
Db 79 SYFNAGLTM-----NYCRNFDADKSPWCY 102

RESULT 30

PLBO  
plasmin (EC 3.4.21.7) precursor - bovine  
N;Alternate names: plasminogen  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 30-Sep-1997 #sequence\_revision 28-Apr-1995 #text\_change 18-Jun-1999  
C;Accession: S45046; A25835; I45961; S03736  
R;Berglund, L.; Andersen, M.D.; Petersen, T.E.  
submitted to the EMBL Data Library, May 1994  
A;Description: Cloning and characterization of the bovine plasminogen cDNA.  
A;Reference number: S45046  
A;Accession: S45046  
A;Molecule type: mRNA  
A;Residues: 1-812 <BBR>  
A;Cross-references: EMBL:X79402; NID:G494962; PIDN:CAA55939.1; PID:G494963  
A;Experimental source: liver

A>Note: it is uncertain whether Met-1 or Met-8 is the initiator  
R;Schaller, J.; Moser, P.W.; Dannegger-Muller, G.A.K.; Rosselet, S.J.; Kampfner, U.; Ri:  
Eur. J. Biochem. 149, 267-278, 1985

A>Title: Complete amino acid sequence of bovine plasminogen. Comparison with human pla:  
A;Reference number: A25835; MUID:85203906; PMID:3846532

A;Accession: A25835

A:Molecule type: Protein

A;Residues: 27-334, D' ,336-515, H', 517-554, L', 556-812 <SCH>

R;Malinowski, R.; Sadler, J.E.; Davie, E.W.  
Biochemistry 23, 4243-4250, 1984

A>Title: Characterization of a complementary deoxyribonucleic acid coding for human an:  
A;Reference number: I45961; MUID:85023311; PMID:6148961

A;Accession: I45961

A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A;Residues: 706-743, R', 745-812 <MAL>

A;Cross-references: GB:X02935; NID:glG35S1; PIDN:AAA30714.1; PID:glG35S2

R;Brundisholz, R.A.; Lerch, P.G.; Schallier, J.; Rickli, B.E.; Lergier, W.; Manneberg, M.:  
Eur. J. Biochem. 114, 465-470, 1981

A>Title: Comparison of the primary structure of the N-terminal CNB fragments of human

A;Reference number: S03735; MUID:81212097; PMID:7238497

A;Accession: S03736

A:Molecule type: protein

A;Residues: 27-83 <BRU>

C>Description:

A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a  
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen a

A;Pathway: fibrinolysis

C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolog

C;Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasm

F;1-26/Domain: signal sequence #status predicted <SIG>  
F;8-103/Domain: plasminogen-related protein precursor homology <PLPH>  
F;27-812/Product: plasminogen #status experimental <PRO>  
F;27-103/Domain: activation peptide #status experimental <APT>  
F;104-583 /Product: plasmin #status experimental <MAT>  
F;104-583/Domain: plasmin chain A #status experimental <ACH>  
F;104-188/Domain: plasmin homology <KRI>  
F;152-269/Domain: kringle homology <KR1>  
F;282-359/Domain: kringle homology <KR2>  
F;384-461/Domain: kringle homology <KR4>  
F;485-564/Domain: kringle homology <KR5>  
F;584-812/Domain: plasmin chain B #status experimental <BCH>  
F;584-805/Domain: trypsin homology <TRY>  
F;58-80, 60-68, 110-188, 131-171, 159-183, 192-269, 195-323, 213-252, 241-264, 282-359, 303-342,  
bonds: #status predicted

F;315/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F;365/Binding site: carbohydrate (Ser) (covalent) #status experimental  
P;624, 667, 762/Active site: His, Asp, Ser #status predicted

	Query Match	Best Local Similarity	Matches	Score	ID	Pred.	No.	Gaps
		19.7%;	Matched 47;	Conservative 47;	Indels 54;	Matches 53;	Length 812;	DB 1;
		28.3%;	Matched 47;	Conservative 47;	Indels 54;	Matches 53;	Length 812;	DB 1;
Qy	5 HQVPSNDCLN-----GGTCVSNKYFSNIHW--CNCPRKKFGQHCEIDKSRT 49	: : :	:	:	:	:	:	-CESSPLST 366
Dd	315 NRTPENFPCKLENCYCRNPNGERAPWCYTIN--SEVRWEYCITPS-----	:	:	:	:	:	:	-CESSPLST 366
Qy	50 -----CYENGHFHYRKASTDTMGRCLPNWSATVLQOTTHAH---RS 89	:	:	:	:	:	:	--RS 89
Dd	367 ERMDVPVPPEQTPTPDQCCHYGNGOSYRGTSITTIGRCQSWS-----MTPIRHLLKTPE 421	:	:	:	:	:	:	----MTPIRHLLKTPE 421
Qy	90 DALQLGLGHKNYCRNPDRRRPCWCVVGKLKPLV--QECSVHDCAAD 133	:	:	:	:	:	:	----
Dd	422 NYFNAGL-TWNVCNPPADSKFCWT---TDPFRWFECNLKCSE 463	:	:	:	:	:	:	---

Search completed: December 3, 2003, 14:44:15  
Job time : 6.98276 secs

Search completed: December 3, 2003, 14:44:15  
Job time : 6.98276 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:53 ; Search time 4.33777 Seconds  
(without alignments)  
1463.563 Million cell updates/sec

Title: US-09-880-503-4  
Perfect score: 793  
Sequence: 1 SNELHQVPSNDCINGTCV.....QVGLKPLVQBMVHDCADGK 135

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	793	100.0	431	1	UROK_HUMAN
2	713.5	90.0	433	1	UROK_PAPCY
3	608	76.7	442	1	UROK_PIG
4	579	73.0	433	1	UROK_BOVIN
5	571	72.0	433	1	UROK_RAT
6	548	69.1	433	1	UROK_MOUSE
7	335.5	42.3	477	1	URT2_DESRO
8	334.5	42.2	434	1	UROK_CHICK
9	328.5	41.4	431	1	URT2_DESRO
10	328.5	41.4	562	1	TPA_HUMAN
11	321.5	40.5	559	1	TPA_RAT
12	315.5	39.8	559	1	TPA_MOUSE
13	310.5	39.2	477	1	URT1_DESRO
14	300.5	37.9	566	1	TPA_BOVIN
15	280.5	32.8	655	1	HGF_MOUSE
16	257	32.4	603	1	FA12_CAVPO
17	250.5	31.6	653	1	HGF_MOUSE
18	233	29.4	615	1	FA12_HUMAN
19	223	28.1	394	1	URT2_DESRO
20	217	27.4	593	1	FA12_BOVIN
21	169	21.3	685	1	ROR1_DROME
22	162.5	20.5	4548	1	APCA_HUMAN
23	161	20.3	473	1	KRM1_RAT
24	161	20.3	1420	1	APCA_MACMU
25	159	20.1	452	1	KRM1_XENLA
26	159	20.1	473	1	KRM1_MOUSE
27	158.5	20.0	937	1	ROR1_HUMAN
28	158.5	20.0	937	1	ROR1_MOUSE
29	156	19.7	475	1	KRM1_HUMAN
30	156	19.7	812	1	PLMN_BOVIN
31	155	19.5	790	1	PLMN_PIG
32	154	19.4	810	1	PLMN_HUMAN
33	153.5	19.4	810	1	PLMN_ERIEU

## RESULT 1

ID	UROK_HUMAN	STANDARD	PRT	431 AA
AC	P00749; Q15844; Q16618; Q969W6;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	20-MAR-1987 (Rel. 04, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)			
DE	(U-plasminogen activator).			
GN	PLAU.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85215647; PubMed=2987867;			
RA	Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;			
RT	"The human urokinase-plasminogen activator gene and its promoter.;"			
RL	Nucleic Acids Res. 13:2759-2771(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,			
RT	Steffens G.J., Heyneker H.L.;			
RL	"Cloning and expression of the gene for pro-urokinase in Escherichia coli.;"			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86056954; PubMed=2415429;			
RA	Nagai M., Hitamatsu R., Kaneda T., Hayasuke N., Atimura H.,			
RT	"Molecular cloning of cDNA coding for human prepro-urokinase.;"			
RL	Gene 36:183-188(1985).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85203359; PubMed=3888571;			
RA	Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,			
RT	van Elsen A., Herzog A., Bollen A.;			
RL	"Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA.;"			
RN	DNA 4:139-146(1985).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RA	Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,			
RT	Nickerson D.A.;			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RT	TISSUE=Lung;			
RX	MEDLINE=22386257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RT	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RL	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			

P12545 macaca mula  
Q01177 rattus norv  
Q09CWO homo sapien  
Q01974 homo sapien  
Q92138 mus musculu  
P20918 mus musculu  
Q8k1s7 mus musculu  
P26927 homo sapien  
P14210 homo sapien  
P80009 canis famil  
P26928 mus musculu  
Q08048 mus musculu

## ALIGNMENTS

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Faley J., Helton E., Kettenan M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [7]  
RP SEQUENCE OF 66-431 FROM N.A.  
RX MEDLINE=84272706; PubMed=6589620;  
RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.;  
RT "Identification and primary sequence of an unspliced human urokinase  
RT poly(A)+ RNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).  
RN [8]  
RP SEQUENCE OF 21-177.  
RX MEDLINE=83055084; PubMed=6754569;  
RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,  
RA Flohe L.;  
RT "The primary structure of high molecular mass urokinase from human  
RT urine. The complete amino acid sequence of the A chain.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).  
RN [9]  
RP SEQUENCE OF 156-176 AND 179-224.  
RX MEDLINE=83003608; PubMed=6749491;  
RA Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W.,  
RA Studer R.O.;  
RT "Human low-molecular-weight urinary urokinase. Partial  
RT characterization and preliminary sequence data of the two polypeptide  
RT chains.";  
RL Eur. J. Biochem. 125:251-257(1982).  
RN [10]  
RP SEQUENCE OF 158-410.  
RX MEDLINE=83055099; PubMed=6754572;  
RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;  
RT "The complete amino acid sequence of low molecular mass urokinase  
RT from human urine.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RX MEDLINE=96000858; PubMed=8591045;  
RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,  
RA Dobson C.M., Stuart D.I., Jones E.Y.;  
RT "The crystal structure of the catalytic domain of human  
RT urokinase-type plasminogen activator.";  
RL Structure 3:681-691(1995).  
RN [12]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.  
RX MEDLINE=20266327; PubMed=10805774;  
RA Spel S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,  
RA Bode W., Magdolen V., Huber R., Moroder L.;  
RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly  
RT selective inhibitors of human urokinase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).  
RN [13]  
RP STRUCTURE BY NMR.  
RX MEDLINE=89127526; PubMed=2536903;  
RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;  
RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-  
RT dimensional NMR.";  
RL Nature 337:579-582(1989).  
RN [14]  
RP STRUCTURE BY NMR OF 67-155.  
RX MEDLINE=93003110; PubMed=1377118;  
RA Li X., Smith R.A.G., Dobson C.M.;

RT "Sequential 1H NMR assignments and secondary structure of the kringle  
RT domain from urokinase.";  
RL Biochemistry 31:9562-9571(1992).  
RN [15]  
RP STRUCTURE BY NMR OF 67-155.  
RX MEDLINE=94149701; PubMed=8107091;  
RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;  
RT "Solution structure of the kringle domain from urokinase-type  
RT plasminogen activator.";  
RL J. Mol. Biol. 235:1548-1559(1994).  
RN [16]  
RP VARIANT LEU-141.  
RX MEDLINE=96186279; PubMed=8652631;  
RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,  
RA Sawasaki Y., Hanada K.;  
RT "Characterization of single chain urokinase-type plasminogen  
RT activator with a novel amino-acid substitution in the kringle  
RT structure.";  
RL Biochim. Biophys. Acta 1293:83-89(1996).  
RN [17]  
RP VARIANT LEU-141.  
RX MEDLINE=97218551; PubMed=9065988;  
RA Conne B., Berczy M., Belin D.;  
RT "Detection of polymorphisms in the human urokinase-type plasminogen  
RT activator gene.";  
RL Thromb. Haemost. 77:434-435(1997).  
RN [18]  
RP ERRATUM.  
RA Conne B., Berczy M., Belin D.;  
RL Thromb. Haemost. 78:973-973(1997).  
RN [19]  
RP VARIANT LEU-141.  
RX MEDLINE=97337920; PubMed=9194591;  
RA Turkmen B., Schmitt M., Schmaifeldt B., Trommler P., Hell W.,  
RA Creutzburg S., Graeff H., Magdolen V.;  
RT "Mutational analysis of the genes encoding urokinase-type plasminogen  
RT activator (upa) and its inhibitor PAI-1 in advanced ovarian cancer.";  
RL Electrophoresis 18:686-689(1997).  
CC -!- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR  
CC THERAPY OF THROMBOLYTIC DISORDERS.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS  
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A  
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW  
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.  
CC -!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used  
CC in Pulmonary Embolism (PE) to initiate fibrinolysis.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X02419; CAA26268.1; -  
DR EMBL; M15476; AAA61253.1; -  
DR EMBL; D00244; BAA00175.1; -  
DR EMBL; D11143; BAA01919.1; -  
DR EMBL; X02760; CAA26535.1; -  
DR EMBL; AF377330; AAK53822.1; -  
DR EMBL; BC013575; AAH13575.1; -  
DR EMBL; K03226; AAC97138.1; -  
DR EMBL; K02286; AAA61452.1; -  
DR EMBL; A21571; CAA01559.1; -  
DR EMBL; A18397; CAA01390.1; -  
DR PIR; A00931; UKHU.  
DR PDB; 1KDU; 31-OCT-93.

```
Query Match      100.0%; Score 793; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.6e-70;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLQQTYHAHRSDALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLQQTYHAHRSDALQGLGKHNYCRNPDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
DB 141 PLVQECMVHDCADGK 155

RESULT 2
UROK_PAPCY STANDARD; PRT; 433 AA.
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
   plasminogen activator.";
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
   plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
   OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
   LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
   MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC -----
DR ENBL; X51935; CAA36200.1; -.
DR PIR; S14687; UKBAY.
DR HSSP; P00749; ILMW.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
```

```
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 26 62 EGF-LIKE.
FT DOMAIN 69 150 KRINGLE.
FT DOMAIN 151 177 CONNECTING PEPTIDE.
FT DOMAIN 178 433 SERINE PROTEASE.
FT DISULFID 30 38 BY SIMILARITY.
FT DISULFID 32 50 BY SIMILARITY.
FT DISULFID 52 61 BY SIMILARITY.
FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
FT DISULFID 208 224 BY SIMILARITY.
FT DISULFID 216 287 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match      90.0%; Score 713.5; DB 1; Length 433;
Best Local Similarity 91.9%; Pred. No. 1.5e-62;
Matches 124; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SREL-QVPSDCGLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 79
QY 61 KASTDTMGRPCLPWNSATVLQQTYHAHRSDALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 80 KASTDTMGRSCLAWSATVLQQTYHAHRSDALQGLGKHNYCRNPDRRRPWCYVQVGLK 139
QY 121 PLVQECMVHDCADGK 135
DB 140 QRVECMVHNCADGK 154

RESULT 3
UROK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
```



```
RL Nucleic Acids Res. 12:9525-9541 (1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.;
CC Submitted (DEC-1986) to the PIR data bank.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01648; CAA25806.1; -.
DR EMBL; X02724; CAA26511.1; -.
DR PIR; A00932; UKFG.
DR HSP; P00749; IKDU.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Plasminogen activation; Hydrolase; 1.
DR Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 442
FT CHAIN 21 188
FT CHAIN 190 442
FT DOMAIN 29 65
FT DOMAIN 72 153
FT DOMAIN 154 189
FT DOMAIN 190 442
FT CARBOHYD 152 152
FT DISULFID 33 41
FT DISULFID 35 53
FT DISULFID 55 64
FT DISULFID 179 310
FT DISULFID 220 236
FT DISULFID 228 299
FT DISULFID 324 393
FT DISULFID 356 372
FT DISULFID 383 411
FT ACT_SITE 235 235
FT ACT_SITE 286 286
FT ACT_SITE 387 387
FT ACT_SITE 241 241
FT CONFLICT 242 242
FT CONFLICT 288 288
FT CONFLICT 288 288
SQ SEQUENCE 442 AA; 49116 MW; EE32FCFE501321EE CRC64;

Query Match 76.7%; Score 608; DB 1; Length 442;
Best Local Similarity 78.1%; Pred. No. 3.2e-52;
Matches 107; Conservative 11; Mismatches 17; Indels 2; Gaps 1;

Nucleic Acids Res. 12:9525-9541 (1984).
[2]
REVISION TO 241.
Nagamine Y.;
Submitted (DEC-1986) to the PIR data bank.
-!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
-!- SIMILARITY: Contains 1 kringle domain.
-!- SIMILARITY: Contains 1 EGF-like domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X01648; CAA25806.1; -.
EMBL; X02724; CAA26511.1; -.
PIR; A00932; UKFG.
HSP; P00749; IKDU.
MEROPS; S01.231; -.
InterPro; IPR001314; Chymotrypsin.
InterPro; IPR006209; EGF-like.
InterPro; IPR000001; Kringle.
InterPro; IPR001254; Ser.protease_Try.
Pfam; PF00051; kringle; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
PRODOM; PD000395; Kringle; 1.
SMART; SM00130; KR; 1.
SMART; SM00020; Tryp_SPC; 1.
PROSITE; PS00022; EGF_1; 1.
PROSITE; PS01186; EGF_2; FALSE_NEG.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS00070; KRINGLE_2; 1.
PROSITE; PS02040; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
Plasminogen activation; Hydrolase; 1.
Kringle; EGF-like domain; Zymogen; Signal.
SIGNAL 1 20
CHAIN 21 442
CHAIN 21 188
CHAIN 190 442
DOMAIN 29 65
DOMAIN 72 153
DOMAIN 154 189
DOMAIN 190 442
CARBOHYD 152 152
DISULFID 33 41
DISULFID 35 53
DISULFID 55 64
DISULFID 179 310
DISULFID 220 236
DISULFID 228 299
DISULFID 324 393
DISULFID 356 372
DISULFID 383 411
ACT_SITE 235 235
ACT_SITE 286 286
ACT_SITE 387 387
ACT_SITE 241 241
CONFLICT 242 242
CONFLICT 288 288
CONFLICT 288 288
SEQUENCE 442 AA; 49116 MW; EE32FCFE501321EE CRC64;

Query Match 76.7%; Score 608; DB 1; Length 442;
Best Local Similarity 78.1%; Pred. No. 3.2e-52;
Matches 107; Conservative 11; Mismatches 17; Indels 2; Gaps 1;
```





OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;  
OC Desmodontinae; Desmodus.  
ON NCBI\_TaxID=9430;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Salivary gland;  
RX MEDLINE=92039036; PubMed=1937019;  
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
RA Alagon A., Donner P., Schleuning W.D.;  
RT "The plasminogen activator family from the salivary gland of the  
RT vampire bat Desmodus rotundus: cloning and expression.";  
RL Gene 105:229-237(1991).  
RN [2]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Salivary gland;  
RX MEDLINE=90036867; PubMed=2509450;  
RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,  
RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;  
RT "Isolation, characterization, and cDNA cloning of a vampire bat  
RT salivary plasminogen activator.";  
RL J. Biol. Chem. 264:17947-17952(1989).  
RN [3]  
RP CHARACTERIZATION.  
RX MEDLINE=93393059; PubMed=1309059;  
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,  
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,  
RA Donner P.;  
RT "Plasminogen activators from the saliva of Desmodus rotundus (common  
RT vampire bat): unique fibrin specificity.";  
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).  
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS  
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC  
CC AGENT.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN  
CC THE PRESENCE OF FIBRIN I.  
CC -!- SUBUNIT: Monomer.  
CC -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,  
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED  
CC STIMULATION OF ACTIVITY.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; MG3988; AAA31593.1; -;  
DR EMBL; J05082; AAA31596.1; -;  
DR PIR; JS0598; JS0598.  
DR HSSP; P98119; 1A51.  
DR MEROPS; S01.232; -;  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR000083; Fibinctn1.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF00039; fni; 1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00181; EGF; 1.  
DR SMART; SM00058; FNI; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation: Hydrolase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Signal; Multigene family.  
FT SIGNAL 1 36  
FT CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 2.  
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.  
FT DOMAIN 83 121 EGF-LIKE.  
FT DOMAIN 128 209 KRINGLE.  
FT DOMAIN 225 477 SERINE PROTEASE.  
FT ACT\_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 321 321 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 428 428 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 42 72 BY SIMILARITY.  
FT DISULFID 70 79 BY SIMILARITY.  
FT DISULFID 87 98 BY SIMILARITY.  
FT DISULFID 92 109 BY SIMILARITY.  
FT DISULFID 111 120 BY SIMILARITY.  
FT DISULFID 128 209 BY SIMILARITY.  
FT DISULFID 149 191 BY SIMILARITY.  
FT DISULFID 180 204 BY SIMILARITY.  
FT DISULFID 214 345 BY SIMILARITY.  
FT DISULFID 257 273 BY SIMILARITY.  
FT DISULFID 265 334 BY SIMILARITY.  
FT DISULFID 359 434 BY SIMILARITY.  
FT DISULFID 391 407 BY SIMILARITY.  
FT DISULFID 424 452 BY SIMILARITY.  
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 403 403 N -> K (IN REF. 2).  
FT CONFLICT 417 417 Y -> H (IN REF. 2).  
FT CONFLICT 435 435 M -> R (IN REF. 2).  
SQ SEQUENCE 477 AA; 53719 MW; 17486555C085077C CRC64;  
  
Query Match 42.3%; Score 335.5; DB 1; Length 477;  
Best Local Similarity 46.3%; Pred. No. 1.4e-25;  
Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;  
  
QY 3 ELHQVP-----SNCDCLNGGTCVSNKYFSNIHWNCNPKFGQHCIEDKSKTCYEGNGHFY 58  
Db 78 QCHTVPKSCSELRCFNGGTCWQAASFDF-VCCPRGYTGKQCEVDTHATCYRQGVTV 136  
QY 59 RGKASTDTWGRPLCPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDNRERPCVQVG 118  
Db 137 RGTWTSSEGAQCINNWNLLTRTYNGRRSDAITLGLGNHNYCRNPDNNSKPCVVIKA 196  
QY 119 LKPLVQECMVHDC 132  
Db 197 SKFILEFCSPVCS 210  
  
RESULT 8  
UROK\_CHICK  
ID UROK\_CHICK STANDARD; PRT; 434 AA.  
AC P15120;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
DE (U-plasminogen activator).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

Gallus.  
NCBI\_TaxID=9031;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=90110185; PubMed=2295632;  
Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;  
"the chicken urokinase-type plasminogen activator gene.";  
J. Biol. Chem. 265:1339-1344 [1990].  
CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
plasminogen to form plasmin.  
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -|- SIMILARITY: Contains 1 kringle domain.  
CC -|- SIMILARITY: Contains 1 EGF-like domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC  
CC EMBL; J05187; AAA49131.1; -;  
CC EMBL; J05188; AAA49130.1; -;  
CC PIR; A35005; A35005.  
CC HSP; P00763; 1DPO.  
CC MEROPS; S01-231; -;  
CC InterPro; IPR001314; Chymotrypsin.  
CC InterPro; IPR006209; EGF-like.  
CC InterPro; IPR006210; EGF.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR001254; Ser.protease\_Try.  
CC Pfam; PF00051; kringle; 1.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PRINTS; PR00018; KRINGLE.  
CC ProDom; PD000395; Kringle; 1.  
CC SMART; SM00181; EGF; 1.  
CC SMART; SM00130; KR; 1.  
CC SMART; SM00020; Tryp\_Spc; 1.  
CC PROSITE; PS00022; EGF\_1; 1.  
CC PROSITE; PS01186; EGF\_2; 1.  
CC PROSITE; PS00021; KRINGLE\_1; 1.  
CC PROSITE; PS00070; KRINGLE\_2; 1.  
CC PROSITE; PS02040; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
CC Kringle; EGF-like domain; Signal; Zymogen.  
CC SIGNAL 1 20  
CC CHAIN 21 434  
CC CHAIN 21 171  
CC CHAIN 173 434  
CC CHAIN 36 72  
CC CHAIN 79 158  
CC DOMAIN 159 172  
CC DOMAIN 173 434  
CC DISULFID 40 48  
CC DISULFID 42 60  
CC DISULFID 62 71  
CC DISULFID 162 296  
CC DISULFID 202 218  
CC DISULFID 210 285  
CC DISULFID 310 379  
CC DISULFID 342 358  
CC DISULFID 369 397  
CC ACT\_SITE 217 217  
CC ACT\_SITE 272 272  
CC ACT\_SITE 373 373  
CC CARBOHYD 228 228  
CC SEQUENCE 434 AA; 49400 MW; BD81048DD666A55 CRC64;  
42.2%; Score 334.5; DB 1; Length 434;  
Query Match

Query Match

```

Best Local Similarity 54.2%; Pred: No. 1.6e-25;
Matches 58; Conservative 14; Mismatches 30; Indels 5; Gaps 2;

Qy 11 CDCLNGGTCVSNKYFSNIHWCNPKFGGQHCIDSKTKCYEGNHGPHYRGKASTDTWGRP 70
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 40 CQCLNGGTCITYRFFSQIKRCLPEGYGGLHCEIDTNSICSYCSGDEYRCMAEDP---G 95

Qy 71 CLPWSNATVLQ-QTYHAHRSDALQLGLGKHNYCRNPENRRRPPWCYVQ 116
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 96 CLYWDHPSVIRWGDYHADLKNALQLGLGKHNYCRNPENGRSRPWCYTK 142

RESULT 9
URTB DESRO
ID_URTB_DESRO STANDARD; PRT; 431 AA.
AC P98121;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
DE beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC --!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC --!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC --!- SUBUNIT: Monomer.
CC --!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC --!- SIMILARITY: Contains 1 kringle domain.
CC --!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63989; AAA31594.1; -.
CC FIR; JS0599; JS0599.
CC DR HSRP; P98119; IASI.
CC DR MEROPS; S01.239; -.
CC DR InterPro; IPR001314; Chymotrypsin.
CC DR InterPro; IPR006209; EGF like.
CC DR InterPro; IPR006210; IEGF.
CC DR InterPro; IPR000001; Kringle.
CC DR InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 1.
CC DR Pfam; PF00051; Kringle; 1.

```











Db 74 QCHSVFVRSCBPRFCNGGTCQALYFSDF-VQCQPDGFVKRCIDIDTRATCFEEQITY 132  
QY 59 RKASTDVTCGRCLPNSATVLTQOYHARSALQGLGKHNYCRPNDRPWCYVQVG 118  
Db 133 RGTWTAESGAECINWSSVLKPYNARRPNAIKLGLGNHNYCRPNDRDLKPCWCVFKA 192  
QY 119 LKPLVQECMVHDCADGK 135  
Db 193 GXYTFECSTPACPKGK 209

RESULT 13  
URTI DESRO  
AC P98119: STANDARD; PRT; 477 AA.  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA  
alpha-1).  
OS Desmodus rotundus (Vampire bat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;  
OC Desmodontinae; Desmodus.  
OX NCBI\_TaxID=9430;  
RN [1] SEQUENCE FROM N.A.  
RP TISSUE=Salivary gland;  
RX MEDLINE=92039036; PubMed=1937019;  
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
Alagon A., Donner P., Schleuning W.D.;  
RT "The plasminogen activator family from the salivary gland of the  
vampire bat Desmodus rotundus: cloning and expression.";  
RL Gene 105:229-237(1991).  
RN [2]  
RP CHARACTERIZATION.  
RX MEDLINE=93393059; PubMed=1309059;  
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,  
Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,  
Donner P.;  
RT "Plasminogen activators from the saliva of Desmodus rotundus (common  
vampire bat): unique fibrin specificity.";  
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
RC TISSUE=Salivary gland;  
RX MEDLINE=98022741; PubMed=9354616;  
RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,  
Schleuning W.D., Bode W.;  
RT "Catalytic domain structure of vampire bat plasminogen activator: a  
molecular paradigm for proteolysis without activation cleavage.";  
RL Biochemistry 36:13483-13493(1997).  
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS  
EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. POTENT THROMBOLYTIC AGENT.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
plasminogen to form plasmin.  
CC -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN  
THE PRESENCE OF FIBRIN I.  
CC -!- SUBUNIT: Monomer.  
CC -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,  
AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED  
STIMULATION OF ACTIVITY.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----  
CC EMBL; M63987; AAA31591.1; -.  
DR EMBL; M63986; AAA31592.1; -.  
DR PIR; JS0597; JS0597.  
DR PDB; 1A51; 23-MAR-99.  
DR MEROPS; S01.232; -.  
DR GlycoSuiteDB; P98119; -.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR000083; Fibrinctn1.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF00039; fnl; 1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PRO0732; CHYMOTRYPSIN.  
DR PRINTS; PRO0018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00058; FNI; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.  
FT SIGNAL; 1 36  
FT CHAIN; 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 1.  
FT DOMAIN; 40 82 FIBRONECTIN TYPE-I.  
FT DOMAIN; 83 121 EGF-LIKE.  
FT DOMAIN; 128 209 KRINGLE.  
FT DOMAIN; 225 477 SERINE PROTEASE.  
FT ACT\_SITE; 272 272 CHARGE RELAY SYSTEM.  
FT ACT\_SITE; 321 321 CHARGE RELAY SYSTEM.  
FT ACT\_SITE; 428 428 CHARGE RELAY SYSTEM.  
FT DISULFID; 42 72 BY SIMILARITY.  
FT DISULFID; 70 79 BY SIMILARITY.  
FT DISULFID; 87 98 BY SIMILARITY.  
FT DISULFID; 92 109 BY SIMILARITY.  
FT DISULFID; 111 120 BY SIMILARITY.  
FT DISULFID; 128 209 BY SIMILARITY.  
FT DISULFID; 149 191 BY SIMILARITY.  
FT DISULFID; 180 204 BY SIMILARITY.  
FT DISULFID; 214 345 BY SIMILARITY.  
FT DISULFID; 257 273 BY SIMILARITY.  
FT DISULFID; 265 334 BY SIMILARITY.  
FT DISULFID; 359 434 BY SIMILARITY.  
FT DISULFID; 391 407 BY SIMILARITY.  
FT DISULFID; 424 452 BY SIMILARITY.  
FT CARBOHYD; 153 153 N-LINKED (GLCNAC. . .).  
FT CARBOHYD; 398 398 /FTID=CAR\_000027.  
FT CARBOHYD; 214 215 N-LINKED (GLCNAC. . .).  
FT CARBOHYD; 223 224 /FTID=CAR\_000028.  
FT TURN; 226 227  
FT STRAND; 230 231  
FT STRAND; 234 236  
FT TURN; 238 239  
FT STRAND; 240 245  
FT STRAND; 254 263  
FT TURN; 264 265  
FT STRAND; 266 269  
FT HELIX; 271 273

```

FT TURN 280 282
FT STRAND 284 287
FT TURN 287 298
FT STRAND 297 309
FT TURN 300 312
FT STRAND 311 317
FT TURN 315 320
FT STRAND 319 328
FT TURN 323 338
FT STRAND 338 340
FT TURN 339 341
FT STRAND 341 345
FT TURN 345 350
FT STRAND 349 356
FT TURN 355 363
FT STRAND 363 366
FT STRAND 374 374
FT STRAND 374 374
FT STRAND 379 385
FT TURN 388 390
FT TURN 393 398
FT TURN 403 404
FT STRAND 405 409
FT TURN 425 426
FT TURN 428 429
FT STRAND 431 436
FT TURN 437 438
FT STRAND 439 448
FT TURN 455 456
FT STRAND 459 463
FT TURN 464 467
FT HELIX 468 474
SQ SEQUENCE 477 AA; 53616 MW; AA06FD1739C10E5E CRC64;

Query Match 39.2%; Score 310.5; DB 1; Length 477;
Best Local Similarity 45.5%; Pred. No. 4e-23;
Matches 60; Conservative 14; Mismatches 53; Indels 5; Gaps 3;

QY 5 HQVPSN-CD---CLNGGTQVSNKYFSNIHWCNCPKFGGHCIDKSKTCYEGNGHYRG 60
DB 80 HTVPVNSCSPPRCFNGGTQWQVFSDF-VQCPAGYTGKRCVDTRATCYEGGVYRG 138
QY 61 KASTDTMGRCPLPNSATV1QOTYHAHRSDALQGLGKHNCRPNDRRFPWCYVQVGLK 120
DB 139 TWSTAESRVEICINWSSLLTFRYNGRMPDAFNGLGNHNYCRNPNGAPKPCWCVIKAGK 198
QY 121 PLVQECWHDCA 132
DB 199 FTSESCVVPVCS 210

RESULT 14
TPA_BOVIN STANDARD; PRT; 566 AA.
AC Q28138;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA.";
RL Int. Dairy J. 5:605-617(1995).
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN

```

TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND MANY OTHER PHYSIOPATHOLOGICAL EVENTS.

-!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

-!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE BOND.

-!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.

-!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER ARG-314 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FACTOR XA.

-!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.

-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

-!- SIMILARITY: Contains 1 EGF-like domain.

-!- SIMILARITY: Contains 1 fibronectin type I domain.

-!- SIMILARITY: Contains 2 kringle domains.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X85800; CAA59795.1; -.  
HSSP; P00750; LRTF.  
MEROPS; S01.232; -.  
InterPro; IPR001314; Chymotrypsin.  
InterPro; IPR006209; EGF\_like.  
InterPro; IPR000083; Fibrinctnl.  
InterPro; IPR006210; IEGF.  
InterPro; IPR000001; Kringle.  
InterPro; IPR001254; Ser\_protease\_Try.  
Pfam; PF000008; EGF; 1.  
Pfam; PF000039; fnl; 1.  
Pfam; PF000051; kringle; 2.  
Pfam; PF000089; trypsin; 1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
PRINTS; PR00018; KRINGLE.  
ProDom; PD000395; Kringle; 2.  
SMART; SM00181; EGF; 1.  
SMART; SM00058; FNI; 1.  
SMART; SM00130; KR; 2.  
SMART; SM00020; Tryp\_SPC; 1.  
PROSITE; PS00022; EGF\_1; 1.  
PROSITE; PS01186; EGF\_2; 1.  
PROSITE; PS01253; FIBRONECTIN\_1; 1.  
PROSITE; PS00021; KRINGLE\_1; 1.  
PROSITE; PS00070; KRINGLE\_2; 2.  
PROSITE; PS00240; TRYPSIN\_DOM; 1.  
PROSITE; PS00134; TRYPSIN\_HIS; 1.  
PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
FT SIGNAL 1 21 BY SIMILARITY.  
FT PROPEP 22 33 TISSUE-TYPE PLASMINOGEN ACTIVATOR.  
FT CHAIN 34 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR A  
FT CHAIN 34 314 CHAIN.  
FT CHAIN 315 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR B  
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.  
FT DOMAIN 83 121 EGF-LIKE.  
FT DOMAIN 128 209 KRINGLE 1.  
FT DOMAIN 219 300 KRINGLE 2.  
FT DOMAIN 315 566 SERINE PROTEASE.  
FT ACT\_SITE 361 361 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 410 410 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 517 517 CHARGE RELAY SYSTEM.  
FT DISULFID 42 72 BY SIMILARITY.



FT DISULFID 108 133 BY SIMILARITY.  
 FT DISULFID 122 148 BY SIMILARITY.  
 FT DISULFID 164 175 BY SIMILARITY.  
 FT DISULFID 169 186 BY SIMILARITY.  
 FT DISULFID 188 197 BY SIMILARITY.  
 FT DISULFID 202 230 BY SIMILARITY.  
 FT DISULFID 228 237 BY SIMILARITY.  
 FT DISULFID 245 256 BY SIMILARITY.  
 FT DISULFID 250 267 BY SIMILARITY.  
 FT DISULFID 269 278 BY SIMILARITY.  
 FT DISULFID 286 367 BY SIMILARITY.  
 FT DISULFID 307 349 BY SIMILARITY.  
 FT DISULFID 338 362 BY SIMILARITY.  
 FT DISULFID 394 521 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 432 448 BY SIMILARITY.  
 FT DISULFID 440 510 BY SIMILARITY.  
 FT DISULFID 535 604 BY SIMILARITY.  
 FT DISULFID 567 583 BY SIMILARITY.  
 FT DISULFID 594 622 BY SIMILARITY.  
 FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 492 492 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 546 546 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 644 644 R -> Q (IN REF. 2).  
 SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1B1B862ED7 CRC64;  
 Query Match 32.8%; Score 260.5; DB 1; Length 655;  
 Best Local Similarity 43.7%; Pred. No. 4.1e-18;  
 Matches 52; Conservative 8; Mismatches 42; Indels 17; Gaps 2;  
 QY 5 HQVPSNCDLNGTGVSNKYFSNIHW-----CNCPKKFGGHCHEIDKSKTCYSGNGH 56  
 Db 242 HTACLSPLNGTGC-----HLVATGTVACPPGAGFLNCNIEDERCFLNGT 292  
 QY 57 FYRGKASTDTMGPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNDRRPPCYV 115  
 Db 293 GYRGVASTSASGLSCLAWNSDLLYQLHVDVSGAALLGLGPHAYCRNPNDRRPPCYV 351  
 RESULT 16  
 ID FA12\_CAVPO STANDARD; PRT; 603 AA.  
 AC Q04962;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)  
 DE (HAF) (fragment).  
 GN F12.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.  
 RC TISSUE=Liver;  
 RX MEDLINE=93003367; PubMed=1390917;  
 RA Semba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,  
 RA Kambara T., Okabe H.;  
 RT "Primary structure of guinea-pig Hageman factor: sequence around the  
 RT cleavage site differs from the human molecule.";  
 RL Biochim. Biophys. Acta 1159:113-121(1992).  
 CC -1- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN  
 CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE  
 CC GENERATION OF BRADYKININ AND ANGIOTENSIN.  
 CC -1- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-ile bonds in factor  
 CC VII to form factor VIIa and factor XI to form factor XIA.  
 CC -1- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A  
 CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY  
 CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST  
 CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR  
 CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -1- SIMILARITY: Contains 1 fibronectin type II domain.  
 CC -1- SIMILARITY: Contains 2 EGF-like domains.  
 CC -1- SIMILARITY: Contains 1 kringle domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X68615; CAA48600.1; -.  
 DR PIR; S28941; S28941.  
 DR HSSP; P00763; 1DPO.  
 DR MEROPS; S01.211; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR000742; EGF\_2.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR006209; EGF\_Like.  
 DR InterPro; IPR000083; Fibrinctn1.  
 DR InterPro; IPR000562; FN Type II.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00039; fn1; 1.  
 DR Pfam; PF00040; fn2; 1.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00013; FNTYPEII.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000995; FN Type II; 1.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 2.  
 DR SMART; SM00058; FN1; 1.  
 DR SMART; SM00059; FN2; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; Tryp SPC; 1.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS05070; KRINGLE\_2; 1.  
 DR PROSITE; PS050240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;  
 KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.  
 FT NON TER 1 1  
 FT SIGNAL <1 18  
 FT CHAIN 19 358 ALPHA-FACTOR XIIA HEAVY CHAIN.  
 FT CHAIN 359 603 ALPHA-FACTOR XIIA LIGHT CHAIN.  
 FT DOMAIN 46 87 FIBRONECTIN TYPE-II.  
 FT DOMAIN 93 130 EGF-LIKE 1.  
 FT DOMAIN 132 172 FIBRONECTIN TYPE-I.  
 FT DOMAIN 173 209 EGF-LIKE 2.  
 FT DOMAIN 216 294 KRINGLE.  
 FT DOMAIN 312 342 PRO-RICH.  
 FT DOMAIN 359 603 SERINE PROTEASE.  
 FT ACT SITE 398 398 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 97 109 BY SIMILARITY.  
 FT DISULFID 103 118 BY SIMILARITY.  
 FT DISULFID 120 129 BY SIMILARITY.  
 FT DISULFID 134 162 BY SIMILARITY.  
 FT DISULFID 160 169 BY SIMILARITY.  
 FT DISULFID 177 188 BY SIMILARITY.

```

FT DISULFID 182 197 BY SIMILARITY.
FT DISULFID 199 208 BY SIMILARITY.
FT DISULFID 216 294 BY SIMILARITY.
FT DISULFID 237 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 345 472 BY SIMILARITY.
FT DISULFID 383 399 BY SIMILARITY.
FT DISULFID 391 461 BY SIMILARITY.
FT DISULFID 422 425 BY SIMILARITY.
FT DISULFID 488 557 BY SIMILARITY.
FT DISULFID 520 536 BY SIMILARITY.
FT DISULFID 547 578 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 248 248 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66795 MW; 48DC6B946FB9ED59 CRC64;

Query Match 32.4%; Score 257; DB 1; Length 603;
Best Local Similarity 40.8%; Pred. No. 8.4e-18;
Matches 49; Conservative 17; Mismatches 46; Indels 8; Gaps 3;

QY 13 CLNGGTCVKNKFSNIHWCNCPKFGQHCETDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db 182 CLNGRCLE--VEGHHLDCPCMGYTGPFCDLDTTASCYEGRGVSRYGMARITVSGAKCQ 238
QY 73 PMSATVLQOTYHAHRS-D-ALQLGLGKHNKCRPNDRRPPWCYVQVGLKPLVQECMVHDC 131
Db 239 RWAS-----EATYRNMTAEQALRGHLGHTFCRPNDRPWCPCFVMGNRLSMEYCDLAQC 294

RESULT 17
HGFA_MOUSE
ID HGFA_MOUSE STANDARD; PRT; 653 AA.
AC Q9R098; Q9R0V4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RN SEQUENCE FROM N.A.
RC STRAIN=BALE/c;
RA Itoh H., Kataoka H., Koono H.;
RT "Mouse hepatocyte growth factor activator.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J.,
RA Yang J., Huan Y.;
RT "Activation of HGF by endogenous HGF activator is required for
RT metanephric kidney morphogenesis in vitro.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING
CC IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF099017; AAF02489.1; -.
DR HSSP; P00763; LDPO.
DR MEROPS; S01.228; -.
DR MGD; MGI:1859281; Hgfac.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibnctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KE; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00400; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_SER; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29 BY SIMILARITY.
FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
FT CHAIN 105 145 FIBRONECTIN TYPE-II.
FT DOMAIN 157 195 EGF-LIKE 1.
FT DOMAIN 197 237 FIBRONECTIN TYPE-I.
FT DOMAIN 238 276 EGF-LIKE 2.
FT DOMAIN 283 364 KRINGLE.
FT DOMAIN 406 653 SERINE PROTEASE.
FT ACT_SITE 445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 105 130 BY SIMILARITY.
FT DISULFID 119 145 BY SIMILARITY.
FT DISULFID 161 172 BY SIMILARITY.
FT DISULFID 166 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 199 227 BY SIMILARITY.
FT DISULFID 225 234 BY SIMILARITY.
FT DISULFID 242 253 BY SIMILARITY.
FT DISULFID 247 264 BY SIMILARITY.
FT DISULFID 266 275 BY SIMILARITY.
FT DISULFID 283 364 BY SIMILARITY.
FT DISULFID 304 346 BY SIMILARITY.
FT DISULFID 392 519 INTERCHAIN (BY SIMILARITY).
FT DISULFID 430 446 BY SIMILARITY.

```

```
FT DISULFID 438 508 BY SIMILARITY.
FT DISULFID 533 602 BY SIMILARITY.
FT DISULFID 565 581 BY SIMILARITY.
FT DISULFID 592 620 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 164 164 G -> W (IN REF. 2).
SQ SEQUENCE 653 AA; 70567 MW; 88B4B20255DF7FDC CRC64;

Query Match 31.6%; Score 250.5; DB 1; Length 653;
Best Local Similarity 42.0%; Pred. No. 3.9e-17;
Matches 50; Conservative 11; Mismatches 41; Indels 17; Gaps 2;

QY 5 HQVPSNCDLNGTGVSNKFSNIHW-----CNCPKFGQGHCEIDKSKTCYEGNGH 56
DB 239 HTACLSPCLNGTGC-----HLIVGTGTVCTCPLGYAGFCNVPTERCFLNGT 289
QY 57 FYRGKASTDTMGRCPLPWNBSATVLQOITYHAHRSDALQLGLGKNEYCRNPDRRRPWCYV 115
DB 290 EYRGVASTAASGLSCLAWNSDLIYQELHVDVAAVLLGLGPHAYCENPKDKRPPWCYV 348

RESULT 18
FA12_HUMAN
ID FA12_HUMAN STANDARD; PRT; 615 AA.
AC P00748; P78339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF).
GN FI2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88007593; PubMed=2888762;
RA Cool D.E., McGilivray R.T.A.;
RT "Characterization of the human blood coagulation factor XII gene.
RT Intron/exon gene organization and analysis of the 5'-flanking
RT region.";
RL J. Biol. Chem. 262:13662-13673(1987).
RN [2]
RP SEQUENCE FROM N.A. AND VARIANTS ALA-207; ASP-545 AND HIS-605.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 4-615 FROM N.A.
RX MEDLINE=86176794; PubMed=3754331;
RA Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA Cortese R.;
RT "cDNA sequence coding for human coagulation factor XII (Hageman).";
RL Nucleic Acids Res. 14:3146-3146(1986).
RN [4]
RP SEQUENCE OF 14-615 FROM N.A.
RX MEDLINE=86033830; PubMed=3877053;
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA McGilivray R.T.A.;
RT "Characterization of human blood coagulation factor XII cDNA.
RT Prediction of the primary structure of factor XII and the tertiary
RT structure of beta-factor XIIa.";
RL J. Biol. Chem. 260:13666-13676(1985).
RN [5]
RP SEQUENCE OF 146-615 FROM N.A.
RX MEDLINE=86216049; PubMed=3011063;
RA Que B.G., Davie E.W.;
```

```
RT "Characterization of a cDNA coding for human factor XII (Hageman
RT factor).";
RL Biochemistry 25:1525-1528(1986).
RN [6]
RP SEQUENCE OF 20-379.
RX MEDLINE=85182674; PubMed=3886654;
RA McMullen B.A., Fujikawa K.;
RT "Amino acid sequence of the heavy chain of human alpha-factor XIIa
RT (activated Hageman factor).";
RL J. Biol. Chem. 260:5328-5341(1985).
RN [7]
RP SEQUENCE OF 354-362 AND 373-615.
RX MEDLINE=83291041; PubMed=6604055;
RA Fujikawa K., McMullen B.A.;
RT "Amino acid sequence of human beta-factor XIIa.";
RL J. Biol. Chem. 258:10924-10933(1983).
RN [8]
RP SEQUENCE OF 561-615 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=96133302; PubMed=8528215;
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;
RT "The novel acceptor splice site mutation 11396(G->A) in the factor
RT XII gene causes a truncated transcript in cross-reacting material
RT negative patients.";
RL Hum. Mol. Genet. 4:1235-1237(1995).
RN [9]
RP CARBOHYDRATE-LINKAGE SITE THR-109.
RX MEDLINE=92184750; PubMed=1544894;
RA Harris R.J., Ling V.T., Spellman M.W.;
RT "O-linked fucose is present in the first epidermal growth factor
RT domain of factor XII but not protein C.";
RL J. Biol. Chem. 267:5102-5107(1992).
RN [10]
RP VARIANT WASHINGTON D.C. SER-590.
RX MEDLINE=90046788; PubMed=2510163;
RA Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,
RA Saito H.;
RT "Coagulation factor XII (Hageman factor) Washington D.C.: inactive
RT factor XIIa results from Cys-571->Ser substitution.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).
RN [11]
RP VARIANT LOCARNO PRO-372.
RX MEDLINE=94325559; PubMed=8049433;
RA Hovinga J.K., Schaller J., Stricker H., Willemin W.A., Furlan M.,
RA Laemmle B.;
RT "Coagulation factor XII Locarno: the functional defect is caused by
RT the amino acid substitution Arg-353->Pro leading to loss of a
RT kallikrein cleavage site.";
RL Blood 84:1173-1181(1994).
RN [12]
RP VARIANT TENRI CYS-53.
RX MEDLINE=99290785; PubMed=10361128;
RA Kondo S., Tokunaga F., Kawano S., Oono Y., Kumagai S., Koide T.;
RT "Factor XII Tenri, a novel cross-reacting material negative factor XII
RT deficiency, occurs through a protease-mediated degradation.";
RL Blood 93:4300-4308(1999).
RN [13]
RP THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN
CC -I- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC CATALYTIC ACTIVITY: Cleaves selectively Arg-|-lle bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC -I- PTM: O- AND N-GLYCOSYLATED.
CC -I- DISEASE: DEFECTS IN FI2 DO NOT CAUSE ANY CLINICAL SYMPTOMS. THE
CC SOLE EFFECT IS THAT WHOLE-BLOOD CLOTTING TIME IS PROLONGED.
CC -I- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIa AND THEN TO BETA-FACTOR XIIa. ALPHA-FACTOR
CC XIIa ACTIVATES FACTOR XI TO FACTOR Xla.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -I- SIMILARITY: Contains 2 EGF-like domains.
CC -I- SIMILARITY: Contains 1 fibronectin type I domain.
CC -I- SIMILARITY: Contains 1 fibronectin type II domain.
```

FT	CARBOHYD	249	249	N-LINKED (GLCNAC. . .).
FT	CARBOHYD	299	299	O-LINKED (POTENTIAL).
FT	CARBOHYD	305	305	O-LINKED (POTENTIAL).
FT	CARBOHYD	308	308	O-LINKED (POTENTIAL).
FT	CARBOHYD	328	328	O-LINKED (POTENTIAL).
FT	CARBOHYD	329	329	O-LINKED (POTENTIAL).
FT	CARBOHYD	337	337	O-LINKED (POTENTIAL).
FT	ACT_SITE	412	412	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	461	461	CHARGE RELAY SYSTEM (BY SIMILARITY).

Query Match 29.4%; Score 233; DB 1; Length 615;  
Best Local Similarity 42.3%; Pred.No. 1.9e-15;  
Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;

Qy	13	CLNGGTCVSNKYPNSNIHWCNCPKFGQHQHCEIDKSKTCYEGNGHFYRGKASTDTMGAPCL	72
Db	183	CLHGGRGLE---VEGHRLCHCPVGYTGPFCDDVTDKASYDGRGLSYRGLARTTTLGAPCQ	239
Qy	73	PWNSATVLOQTY-HAHRSDALQLGLGKHNVCNPNDRRRCVY	115
Db	240	PWAS-----EATYRNVTAEQARNWGLGHAFCRNPNDIRPWCVF	279

RESULT 19

URTG	DESRO	
ID	URTG_DESRO	STANDARD; PRT; 394 AA.
AC	P49150;	
DT	01-FEB-1996 (Rel. 33, Created)	
DT	01-FEB-1996 (Rel. 33, Last sequence update)	
DT	28-FEB-2003 (Rel. 41, Last annotation update)	
DE	Salivary plasminogen activator gamma precursor (BC 3.4.21.68) (DSPA gamma).	
DE	Desmodus rotundus (Vampire bat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;	
OC	Desmodontinae; Desmodus.	
ON	NCBI_TaxID=9430;	
OX	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Salivary gland.	
RC	MEDLINE=92039036; PubMed=1937019;	
RA	Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,	
RA	Alagon A., Donner P., Schleuning W.D.;	
RT	"The plasminogen activator family from the salivary gland of the	
RT	vampire bat Desmodus rotundus: cloning and expression.";	
RL	Gene 105:229-237(1991).	
RN	[2]	
RP	CHARACTERIZATION.	
RP	MEDLINE=93393059; PubMed=1309059;	
RA	Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,	
RA	Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,	
RA	Donner P.;	
RT	"Plasminogen activators from the saliva of Desmodus rotundus (common	
RT	vampire bat): unique fibrin specificity.";	
RL	Ann. N.Y. Acad. Sci. 667:395-403(1992).	
CC	-1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS	
CC	EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC	
CC	AGENT.	
CC	-1- CATALYTIC ACTIVITY: Specific cleavage of Arg- -Val bond in	
CC	plasminogen to form plasmin.	
CC	-1- SUBUNIT: Monomer.	
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.	
CC	-1- SIMILARITY: Contains 1 kringle domain.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/</a> or	
CC	send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC	EWBI. M63090. M63150F 1	
DB		



```
DR PIR; JS0600; JS0600.
DR HSSP; P98119; IASI.
DR MEROPS; S01.239; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR SMART; SM00130; KE; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; Signal; Multigene family.
FT SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT DOMAIN 45 126 KRINGLE.
FT DOMAIN 142 394 SERINE PROTEASE.
FT ACT_SITE 189 169 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 45 126 BY SIMILARITY.
FT DISULFID 66 108 BY SIMILARITY.
FT DISULFID 97 121 BY SIMILARITY.
FT DISULFID 131 262 BY SIMILARITY.
FT DISULFID 174 190 BY SIMILARITY.
FT DISULFID 182 251 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.
FT DISULFID 308 324 BY SIMILARITY.
FT DISULFID 341 369 BY SIMILARITY.
FT CARBOHYD 315 315 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 394 AA; 44105 MW; 9CCD6F52F3D81FCD CRC64;

Query Match 28.1%; Score 223; DB 1; Length 394;
Best Local Similarity 44.3%; Pred. No. 1.2e-14;
Matches 39; Conservative 13; Mismatches 36; Indels 0; Gaps 0;

QY 45 DKSCTCYEGHFGYRKASTDTMGRCPLPWSATVLQOYTHAHRSDALQGLGKKNYCRN 104
Db 40 DPAICYKQGVYRTWTSTSGACINWNSLLIRRYNGMPAVKGLGNHNYCRN 99

QY 105 PDNRPRWCYVQGLKPLVQECMVHDC 132
Db 100 PDGASKFCYVTKARKFTSESCSVPVCS 127

RESULT 20
FA12_BOVIN STANDARD; PRT; 593 AA.
AC P98140;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF) (Fragment).
GN F12.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94342782; PubMed=8186251;
RA Shibuya Y., Senba U., Okabe H., Kambara T., Yamamoto T.;
RT "Primary structure of bovine Hageman factor (blood coagulation factor
```

```
RT XII): comparison with human and guinea pig molecules.";
RL Biochim. Biophys. Acta 1206:63-70(1994).
[2]
RN SEQUENCE OF 10-21; 350-364 AND 525-550.
RX MEDLINE=77182112; PubMed=861210;
RA Fujikawa K., Walsh A.K., Davie W.E.;
RT "Isolation and characterization of bovine factor XII (Hageman factor).";
RL Biochemistry 16:2270-2278(1977).
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor VII to form factor VIIa and factor XI to form factor Xla.
CC -!- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).
CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY FACTOR XII TO FORM KALLIKREIN WHICH THEN CLEAVES FACTOR XII FIRST TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR XIIA ACTIVATES FACTOR XI TO FACTOR XIA. BOVINE FACTOR XII IS CLEAVED ONLY TO ALPHA-FACTOR XIIA AS IT LACKS THE TRYPSIN/KALLIKREIN CLEAVAGE SITE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
DR EMBL; S70164; AAB30804.2; -.
DR PIR; S45281; S45281.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.211; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease; Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
```



or send an email to license@isb-sib.ch).

```
CC CC      EMBL; L20297; AAA28860.1; -.
CC CC      EMBL; AJ003628; AAF52885.1; -.
DR DR      EMBL; AJ002908; CAA05743.1; -.
DR DR      PIR; A48289; A48289.
DR DR      HSSP; P11362; IFGX.
DR DR      FlyBase; Fgn0010407; Ror.
DR DR      GO; GO:0016021; C:integral to membrane; NAS.
DR DR      GO; GO:0004713; P:protein tyrosine kinase activity; NAS.
DR DR      GO; GO:0007417; P:central nervous system development; IEP.
DR DR      GO; GO:0006468; P:protein amino acid phosphorylation; NAS.
DR DR      InterPro; IPR000024; Fz domain.
DR DR      InterPro; IPR000001; Kringle.
DR DR      InterPro; IPR000719; Prot_kinase.
DR DR      InterPro; IPR002011; RTKinaseII.
DR DR      InterPro; IPR001245; Tyr_kinase.
DR DR      Pfam; PF00051; kringle; 1.
DR DR      Pfam; PF00069; pkinase; 1.
DR DR      PRINTS; PR00018; KRINGLE
DR DR      PRINTS; PR00109; TYRKINASE.
DR DR      ProDom; PD000395; Kringle; 1.
DR DR      ProDom; PD000001; Prot_kinase; 1.
DR DR      SMART; SM00130; KR; 1.
DR DR      SMART; SM00219; TyrKC; 1.
DR DR      PROSITE; PS50038; FZ; 1.
DR DR      PROSITE; PS00021; KRINGLE_1; 1.
DR DR      PROSITE; PS00021; KRINGLE_2; 1.
DR DR      PROSITE; PS50070; PROTEIN KINASE ATP; 1.
DR DR      PROSITE; PS00107; PROTEIN KINASE DOM; 1.
DR DR      PROSITE; PS50011; PROTEIN KINASE TVR; 1.
DR DR      PROSITE; PS00109; PROTEIN KINASE TVR; 1.
DR DR      PROSITE; PS00239; RECEPTOR TYR KIN II; 1.
DR DR      Transferase; Kinase; Tyrosine-protein kinase; ATP-binding; Receptor;
KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
KW Developmental protein.
FT SIGNAL 1 24
FT CHAIN 25 685
FT DOMAIN 25 317
FT TRANSMEM 318 338
FT DOMAIN 339 685
FT DOMAIN 36 225
FT DOMAIN 236 310
FT DOMAIN 410 677
FT NP_BIND 416 424
FT BINDING 442 442
FT ACT_SITE 539 539
FT MOD_RES 565 565
FT MOD_RES 569 569
FT MOD_RES 570 570
FT MOD_RES 570 570
FT CARBOHYD 45 45
FT CARBOHYD 63 63
FT CARBOHYD 129 129
FT CARBOHYD 144 144
FT CARBOHYD 250 250
FT SEQUENCE 685 AA; 78142 MW; 526162027D5FD7C7 CRC64;
Query Match 21.3%; Score 169; DB 1; Length 685;
Best Local Similarity 28.8%; Pred. No. 3.8e-09;
Matches 38; Conservative 27; Mismatches 35; Indels 32; Gaps 6;
QY 3 ELHQVPSNCDLNGFCVSNKYFSNTHWCNPKFKGQHCEIDKSKTCVEGNCHFRGKA 62
DB 212 DOOKLPQHKDCSLGITI-----EVDKTCYWGDSGYRGVA 249
QY 63 STDTMGRCPLMNSATVLOQTYHAHRSDALQLGLGHKNYCRNPDN-RRRPQVYQVGLXP 121
DB 250 NVSASGKPCLRW--SWLMKEI-----SDPEL-IGQ-NYCRNPGSVENSPWCFVDSRR 300
QY 122 LVQECMVHDCAD 133
```

```
DB 301 ITELCDIPKCAD 312
RESULT 22
APOA_HUMAN
ID APOA_HUMAN STANDARD; PRT; 4548 AA.
AC P08519;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).
GN LPA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=88039109; PubMed=3670400;
RA McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,
RA Fless G.M., Scanu A.M., Lawn R.M.;
RT "cDNA sequence of human apolipoprotein(a) is homologous to
RT plasminogen.";
RL Nature 330:132-137(1987).
RN [2]
SERINE PROTEASE ACTIVITY.
RX MEDLINE=90076123; PubMed=2531657;
RA Salonen E.-M., Jauhainen M., Zardi L., Vaheiri A., Ehnholm C.;
RT "Lipoprotein(a) binds to fibronectin and has serine proteinase
RT activity capable of cleaving it.";
RL EMBO J. 8:4035-4040(1989).
RN [3]
REVIEW.
RX MEDLINE=90049223; PubMed=2530631;
RA Utermann G.;
RT "The mysteries of lipoprotein(a).";
RL Science 246:904-910(1989).
RN [4]
CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
RX MEDLINE=21303595; PubMed=11294842;
RA Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;
RT "Structural elucidation of the N- and O-glycans of human
RT apolipoprotein(a): role of O-glycans in conferring protease
RT resistance.";
RL J. Biol. Chem. 276:22200-22208(2001).
RN [5]
X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
RX MEDLINE=96217891; PubMed=8642595;
RA Mikol V., Lograsso P.V., Boettcher B.R.;
RT "Crystal structures of apolipoprotein(a) kringle IV37 free and
RT complexed with 6-aminohexanoic acid and with p-aminomethylbenzoic
RT acid: existence of novel and expected binding modes.";
RL J. Mol. Biol. 256:751-761(1996).
RN [6]
VARIANT ARG-4193.
RX MEDLINE=95002201; PubMed=7918682;
RA Scanu A.M., Pfaffinger D., Lee J.C., Hinman J.;
RT "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37
RT associated with a lysine binding defect in Lp(a).";
RL Biochim. Biophys. Acta 1227:41-45(1994).
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin.
CC -!- PTM: N- and O-glycosylated. The N-glycans are complex biantennary
CC structures present in either a mono- or disialylated state. The
CC O-glycans are mostly (80%) represented by the monosialylated core
CC type I structure, NeuNacalpha2-3Galbeta1-3GalNAC, with smaller
CC amounts of disialylated and non-sialylated O-glycans also
CC detected.
```





```

RESULT 25
KRM1_XENLA
ID KRM1_XENLA STANDARD; PRT; 452 AA.
AC Q90Y90;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringie-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
[1]
RP SEQUENCE FROM N.A.
RA Nakamura T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringie domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AB070851; BAB4294.1; -.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR000001; Kringie.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringie; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringie; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00321; WSC; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR KX Wnt signaling pathway; Glycoprotein; Kringie;
KW Transmembrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 452
FT DOMAIN 23 369
FT TRANSMEM 370 390
FT DOMAIN 391 452
FT DOMAIN 29 112
FT DOMAIN 118 208
FT DOMAIN 212 319
FT CARBOHYD 43 43
FT CARBOHYD 57 57
FT CARBOHYD 215 215
FT CARBOHYD 253 253
FT CARBOHYD 291 291
FT CARBOHYD 328 328
FT CARBOHYD 344 344
FT CARBOHYD 452 452
SQ SEQUENCE 452 AA; 50188 MW; ED24BCD1AF4564E2 CRC64;

Query Match 20.1%; Score 159; DB 1; Length 452;
Best Local Similarity 40.8%; Pred No. 2, 4e-08;
Matches 31; Conservative 9; Mismatches 32; Indels 4; Gaps 2;

42 CEIDKSKTCYGNHFGYKASTDTM--GRCLPWSATVLTQTYVHAHRSDALQLGLGKH 99

```

```

Db 22 CSDSPHSECYTVNGADYRGQTQNTSLDGGKPCLFWNE--TFQHPNTLKYPNGEGLGKH 79
QY 100 NYCNPDPNRRPWCYV 115
Db 80 NYCNPDPDGVSPWCYI 95

RESULT 26
KRM1_MOUSE
ID KRM1_MOUSE STANDARD; PRT; 473 AA.
AC Q9NM43;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringie-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.; DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.
RA Nakamura T.;
RL MEDLINE=21167372; PubMed=11267660;
RA Nakamura T.;
RT "Molecular cloning and characterization of Kremen, a novel
RT kringie-containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: In the adult, widely expressed with high
CC levels in heart, lung, kidney, skeletal muscle and testis.
CC -!- DEVELOPMENTAL STAGE: In the embryo, expression is first detected
CC on day 9 and increases up to day 18. Lower levels are found in
CC adult. At 9.5 dpc, expression is localised to the apical
CC ectodermal ridge (AER) of the developing fore- and hindlimb buds,
CC the telencephalon and the first brachial arch. At 10.5 dpc,
CC expression is also observed in the myotome and in sensory tissues
CC such as the nasal pit and optic vesicle.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringie domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AB059617; BAB40968.1; -.
DR HSPSP; P00747; 1CEA.
DR MGD; MGI:1933988; Kremen.
DR GO; GO:0016021; C: integral to membrane; NAS.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR000001; Kringie.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringie; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringie; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR KX Wnt signaling pathway; Signal; Transmembrane; Kringie.

```



```
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1 549 Missing (in isoform Short).
FT SEQUENCE 937 AA; 104312 MW; 0D06945DB29F4773 CRC64;

Query Match 20.0%; Score 158.5; DB 1; Length 937;
Best Local Similarity 29.7%; Pred. No. 5.4e-08;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

QY 2 NELHQP-----SNCDGLNGTGVSNKYFS-----NIHWCN-----PKKFGGQH 41
Db 236 DETSSVFKPRDLRCRDECEILENVLCQTEYIFARSNPMLRLKLPNCEDLPQESPEAAN 295

QY 42 C-----EIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTVHAHRS DAL 92
Db 296 CIRIGIPMADPINKNKHCKYCNSTGVDYRGVTSVTKSGRCQCPWNS-----QYPHTHSFTAL 350

QY 93 QLQ--LGKHNCRPNPNRR--PWCY 114
Db 351 RPEELNGGHSYCRNPGNQKEAPWCF 375

RESULT 28
RORI_MOUSE
ID RORI_MOUSE STANDARD; PRT; 937 AA.
AC Q9Z139;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor
DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 1)
DE (mROR1).
DE ROR1 OR NTRK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=9248426; PubMed=102311392;
RA Oishi I.; Takeuchi S.; Hashimoto R.; Nagabukuro A.; Ueda T.; Liu Z.J.;
RA Hattori I.; Akira S.; Matsuda Y.; Yamamura H.; Otani H.; Minami Y.;
RT "Spatio-temporally regulated expression of receptor tyrosine kinases,
RT mRor1, mRor2, during mouse development: Implications in development
RT and function of the nervous system.";
RL Genes Cells 4:41-56(1999).
CC -!- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHOSE ROLE IS NOT YET
CC CLEAR.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB010383; BAA75480.1; -.
CC HSSP; P00747; ICEA.
CC MGD; MGI:1347520; Ror1.
CC InterPro; IPR000024; Fz domain.
CC InterPro; IPR007110; Ig-like.
```

```
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000335; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00219; TyrKC; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Receptor;
KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
KW Immunoglobulin domain.
FT SIGNAL 1 29
FT CHAIN 30 937
FT TYROSINE-PROTEIN KINASE TRANSMEMBRANE
FT RECEPTOR ROR1.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT IG-LIKE C2-TYPE.
FT FZ.
FT KRINGLE.
FT PROTEIN KINASE.
FT SER/THR-RICH.
FT PRO-RICH.
FT SER/THR-RICH.
FT NP_BIND 479 487
FT BINDING 506 506
FT ACT_SITE 615 615
FT MOD_RES 645 645
FT DISULFID 79 131
FT CARBOHYD 47 47
FT CARBOHYD 66 66
FT CARBOHYD 184 184
FT CARBOHYD 315 315
FT SEQUENCE 937 AA; 104156 MW; D728733E67D1782C CRC64;

Query Match 20.0%; Score 158.5; DB 1; Length 937;
Best Local Similarity 29.7%; Pred. No. 5.4e-08;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

QY 2 NELHQP-----SNCDGLNGTGVSNKYFS-----NIHWCN-----PKKFGGQH 41
Db 236 DETSSVFKPRDLRCRDECEILENVLCQTEYIFARSNPMLRLKLPNCEDLPQESPEAAN 295

QY 42 C-----EIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTVHAHRS DAL 92
Db 296 CIRIGIPMADPINKNKHCKYCNSTGVDYRGVTSVTKSGRCQCPWNS-----QYPHTHSFTAL 350

QY 93 QLQ--LGKHNCRPNPNRR--PWCY 114
Db 351 RPEELNGGHSYCRNPGNQKEAPWCF 375

RESULT 29
KRM1_HUMAN
ID KRM1_HUMAN STANDARD; PRT; 475 AA.
AC Q96M08; Q9BY70; Q9UGS5; Q9UGU1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
```



```

DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringele-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Nakamura T., Nakamura T.
RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Koniya M., Sugiyama T., Irie R.,
RA Okusaki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.
RN "NEDO human cDNA sequencing project."
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagnoley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tanlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Levenshina M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spraggon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley H., Wilmer T.E., Wilming L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.B., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoehima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,
RA Sroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkison P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tilihun Y., Wright H.
RT "The DNA sequence of human chromosome 22."
RL Nature 402:489-495(1999).
CC -1- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).

```

```

CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96MU8-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q96MU8-2; Sequence=VSP_003900;
CC Note=No experimental confirmation available;
CC -1- SIMILARITY: Contains 1 CUB domain.
CC -1- SIMILARITY: Contains 1 kringele domain.
CC -1- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB059618; BAB40969.1; -
CC EMBL; AK056425; BAB71180.1; -
CC EMBL; Z95116; CAB62952.1; -
CC EMBL; AL021393; CAB62959.1; -
CC Genew; HGNC:17550; KREMEN1.
CC GO; GO:0016021; C:integral to membrane; ISS.
CC GO; GO:0005624; C:membrane fraction; TAS.
CC GO; GO:0007154; P:cell communication; TAS.
CC InterPro; IPR000859; CUB domain.
CC InterPro; IPR000001; Kringele.
CC Pfam; PF00431; CUB; 1.
CC Pfam; PF00051; kringele; 1.
CC Pfam; PF01822; WSC; 1.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringele; 1.
CC SMART; SM00042; CUB; 1.
CC SMART; SM00130; KR; 1.
CC PROSITE; PS01180; CUB; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC Wnt signaling pathway; Signal; Transmembrane; Kringele;
CC Alternative splicing.
CC SIGNAL 1 19
CC CHAIN 20 475 POTENTIAL.
CC KREMEN PROTEIN 1.
CC EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 395 415 POTENTIAL.
CC DOMAIN 416 475 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 33 115 KRINGLE.
CC DOMAIN 121 202 WSC.
CC DOMAIN 216 323 CUB.
CC CARBOHYD 47 47 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 219 219 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 335 335 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 347 347 N-LINKED (GLCNAC. .) (POTENTIAL).
CC VARSPLIC 473 475 2).
CC /FTId=VSP_003900.
CC MISSING (IN REF. 1).
CC I -> V (IN REF. 2).
CC SEQUENCE 475 AA; 51898 MW; B7E86FD80F96A0A4 CRC64;
Query Match 19.7%; Score 156; DB 1; Length 475;
Best Local Similarity 44.1%; Pred. No. 4.9e-08;
Matches 30; Conservative 8; Mismatches 26; Indels 4; Gaps 2;
QY 50 CYEGNGHFRGKASTDTM--GRPLPWNATVLQQTTHAHRSDALQLGLGKHYCNPDN 107
Db 34 CFANGADYRGTONWTALQGGKPLFWNE--TFQHPYNTLKYFNGEGGLGKHYCNPDG 91
QY 108 RRRPCWCV 115
|||||

```



Search completed: December 3, 2003, 14:40:07  
Job time : 5.33777 secs

---

GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:23 ; Search time 16.6105 Seconds  
 (without alignments)  
 2097.294 Million cell updates/sec

Title: US-09-880-503-4  
 Perfect score: 793  
 Sequence: 1 SNELHQVPSNCDLNGTGV.....QVGLKPLVQECMVHDCADCK 135

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database: SPTREMBL 23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvivirus:\*
- 16: sp\_bacteriophage:\*
- 17: sp\_archaeop:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	778	98.1	154	Q96SE8	Q96se8 homo sapien
2	602	75.9	433	Q8MIL0	Q8milo oryctolagus
3	602	75.9	433	Q8MHI7	Q8mhy7 oryctolagus
4	550	69.4	157	Q9TVB8	Q9tvb8 bos taurus
5	548	69.1	231	Q8C6L2	Q8c6l2 mus musculus
6	322.5	40.7	516	Q9BU99	Q9bu99 homo sapien
7	315.5	39.8	559	Q91VP2	Q91vp2 mus musculus
8	313	39.5	214	Q9XT70	Q9xt70 oryctolagus
9	310.5	39.2	562	Q8SQ23	Q8sq23 sus scrofa
10	301.5	38.0	564	Q8MKB1	Q8mbk1 oryctolagus
11	264.5	33.4	517	Q8KOD2	Q8kod2 mus musculus
12	257.5	32.5	560	Q14520	Q14520 homo sapien
13	257	32.4	616	Q97507	Q97507 sus scrofa
14	250.5	31.6	653	Q8VCS4	Q8vcs4 mus musculus
15	249	31.4	597	Q35727	Q35727 mus musculus
16	233	29.4	615	Q8I2Z5	Q8i2z5 homo sapien

17	229.5	28.9	395	4	Q9BZW1	Q9bzw1 homo sapien
18	218.5	27.6	90	4	Q8NG20	Q8ng20 homo sapien
19	197.5	24.9	202	13	Q90675	Q90675 gallus gall
20	174	21.9	421	13	Q8AXX3	Q8axx3 xenopus lae
21	170.5	21.5	947	13	Q8AXY6	Q8axy6 gallus gall
22	166	20.9	716	13	Q91691	Q91691 xenopus lae
23	165.5	20.9	359	6	Q8WNR1	Q8wmr1 canis famil
24	165	20.8	812	11	Q9R0W3	Q9r0w3 rattus norv
25	159	20.1	452	13	Q90Y90	Q90y90 xenopus lae
26	158.5	20.0	393	4	Q9BRB6	Q9brb6 homo sapien
27	158.5	20.0	937	11	Q8BNP9	Q8bnp9 mus musculu
28	158.5	20.0	937	11	Q8BG10	Q8bg10 mus musculu
29	158	19.9	454	6	Q46506	Q46506 papio hamad
30	156	19.7	113	4	Q9UIR5	Q9uir5 homo sapien
31	154	19.4	806	6	Q18783	Q18783 macropus eu
32	154	19.4	810	4	Q15146	Q15146 homo sapien
33	153	19.3	113	4	Q9UIR7	Q9uir7 homo sapien
34	152.5	19.2	717	13	P70006	P70006 xenopus lae
35	151.5	19.1	648	4	Q9H1V4	Q9hlv4 homo sapien
36	150.5	19.0	567	4	Q13208	Q13208 homo sapien
37	150.5	19.0	930	13	Q8AV69	Q8av69 xenopus lae
38	149	18.8	801	11	Q8K0Q8	Q8k0q8 mus musculu
39	149	18.8	944	11	Q8C3W2	Q8c3w2 mus musculu
40	149	18.8	944	11	Q8BSP6	Q8bsp6 mus musculu
41	148.5	18.7	381	4	Q8N2J4	Q8n2j4 homo sapien
42	148.5	18.7	399	4	Q96GL8	Q96gl8 homo sapien
43	148.5	18.7	420	4	Q9BTP9	Q9btp9 homo sapien
44	148.5	18.7	424	4	Q8NCW1	Q8ncw1 homo sapien
45	148.5	18.7	462	4	Q8NCW0	Q8ncw0 homo sapien

## ALIGNMENTS

RESULT 1

Q96SE8 PRELIMINARY; PRT; 154 AA.

ID Q96SE8; 01-DEC-2001 (TREMREL. 19, Created)

AC Q96SE8; 01-DEC-2001 (TREMREL. 19, Last sequence update)

DT 01-DEC-2001 (TREMREL. 19, Last sequence update)

DT 01-MAR-2003 (TREMREL. 23, Last annotation update)

DE Urokinase-type plasminogen activator amino-terminal fragment.

GN ATF.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Fu J., Bai X., Ruan C.;

RT "Cloning and expression of the amino-terminal fragment of human urokinase-type plasminogen activator.";

RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Bai X., Fu J., Wang W., Xi X., Ruan C.;

RT "Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator in breast cancer cells results in decreased tumor invasion, growth and angiogenesis.";

RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL: AY029537; AAK38734.1; -.

DR InterPro: IPR006209; EGF like.

DR InterPro: IPR006210; IEGF.

DR InterPro: IPR000001; Kringle.

DR Pfam: PF00051; kringle; 1.

DR PRINTS: PR00018; KRINGLE.

DR ProDom: PD000395; Kringle; 1.

DR SMART: SM00181; EGF; 1.

DR SMART: SM00130; KG; 1.

DR PROSITE: PS00022; EGF 1; 1.

DR PROSITE: PS00021; KRINGLE 1; 1.

DR PROSITE: PS00070; KRINGLE\_2; 1.

```
KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CCF2FCFF505572 CRC64;

Query Match 98.1%; Score 778; DB 4; Length 154;
Best Local Similarity 99.3%; Pred. No. 1.8e-77;
Matches 133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKFGQHCEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKFGQHCEDKSKTCYEGNGHFYRG 80

Qy 61 KASDTMTGRPCLPWNSATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQGLK 120
Db 81 KASDTMTGRPCLPWNSATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQGLK 140

Qy 121 PLVQECMVHDCADG 134
Db 141 LLVQECMVHDCADG 154

RESULT 2
Q8MIL0 PRELIMINARY; PRT; 433 AA.
AC Q8MIL0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22155945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dichek D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
RT development causes arterial constriction and lumen loss, and
RT accelerates lesion growth.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY122285; AAM83187.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR PRINTS; PR0006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65B64F36415549B0 CRC64;

Query Match 75.9%; Score 602; DB 6; Length 433;
Best Local Similarity 76.6%; Pred. No. 1.1e-57;
Matches 105; Conservative 10; Mismatches 20; Indels 2; Gaps 1;

Qy 1 SNELHQV--PSNCDCLNGGTCVSNKYFSNIHWCNPKFGQHCEDKSKTCYEGNGHFY 58
Db 21 SNELHGVSDASNCGLNGGTCVTYKYFSNIWRCNPKFKFQGEHCEDTLTKTCYHGDGHSY 80

Qy 59 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVG 118
Db 81 RGKANTDMDRPPCLAWNSANVLTQTYHAHRPDALQLGLGKHNYCRNPDRHQRPPWCYVQVG 140

Qy 119 LKPLVQECMVHDCADGK 135
Db 141 LKQLIQECKVHDCSSGK 157

RESULT 3
Q8MHY7 PRELIMINARY; PRT; 433 AA.
AC Q8MHY7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029517; AAK40239.1; -.
DR EMBL; AB087224; BAC02685.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match 75.9%; Score 602; DB 6; Length 433;
Best Local Similarity 76.6%; Pred. No. 1.1e-57;
Matches 105; Conservative 10; Mismatches 20; Indels 2; Gaps 1;

Qy 1 SNELHQV--PSNCDCLNGGTCVSNKYFSNIHWCNPKFKFGQHCEDKSKTCYEGNGHFY 58
Db 21 SNELHGVSDASNCGLNGGTCVTYKYFSNIWRCNPKFKFQGEHCEDTLTKTCYHGDGHSY 80

Qy 59 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVG 118
Db 81 RGKANTDMDRPPCLAWNSANVLTQTYHAHRPDALQLGLGKHNYCRNPDRHQRPPWCYVQVG 140

Qy 119 LKPLVQECMVHDCADGK 135
Db 141 LKQLIQECKVHDCSSGK 157
```

```
RESULT 4
Q9TVA8 PRELIMINARY; PRT; 157 AA.
AC Q9TVA8;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21071388; PubMed=11204721;
RA Balcerzak D., Quereguesser L., Dixon W.T., Batacos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
activators and inhibitors in bovine skeletal muscle.";
RL J. Anim. Sci. 79:94-107(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF144761; AAD30301.1; -.
DR HSP; P00749; IURK.
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; Kringle.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kinase; Kringle.
FT NON_TER 1
FT TER 157
FT NCBI_TaxID=9913;
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 69.4%; Score 550; DB 6; Length 157;
Best Local Similarity 75.6%; Pred. No. 1.8e-52;
Matches 93; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 13 CLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFGYRGKASTDTMGRPCL 72
Db 1 CLNGGKCVTYKYFSNIQRCSCKPKFGGQHCIDTSTKYCGNGHSGYRGKANDLSGRPCL 60

QY 73 PWSNATVLQOYTHAHRSDALQLGLGKHNYCRNPDRRPPWCYVQVGLKPLVOECMVHDC 132
Db 61 AWDSPTVLKLMYHAHRSDALQLGLGKHNYCRNPDRRPPWCYVQVGLKPLVOECMVHDC 120

QY 133 DGK 135
Db 121 VGK 123

RESULT 5
Q8C6L2 PRELIMINARY; PRT; 231 AA.
AC Q8C6L2;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;

"Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK054349; BAC35743.1; -.
FT NON_TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25B8980A682737F2 CRC64;

Query Match 69.1%; Score 548; DB 11; Length 231;
Best Local Similarity 71.7%; Pred. No. 4.5e-52;
Matches 91; Conservative 13; Mismatches 23; Indels 0; Gaps 0;

QY 9 SNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFGYRGKASTDTMG 68
Db 30 SNCGGNGGVCSYKYFSRIKRCSCPRFQGHCHIDASKTCYHNGDSYRGKANTDTKG 89

QY 69 RPCLPWSNATVLQOYTHAHRSDALQLGLGKHNYCRNPDRRPPWCYVQVGLKPLVOECMV 128
Db 90 RPCLAWNAPVLQKPYNAHRPDALSLGLGKHNYCRNPDRRPPWCYVQVGLKPLVOECMV 149

QY 129 HDCADGK 135
Db 150 HDCSLSK 156

RESULT 6
Q9BU99 PRELIMINARY; PRT; 516 AA.
AC Q9BU99;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC002795; AAH02795.1; -.
DR HSP; P00750; IASH.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 2.
DR PROSITE; PS00020; TRYD_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
RN [1]
RP SEQUENCE 516 AA; 57370 MW; BA831901FDC96800 CRC64;

Query Match 40.7%; Score 322.5; DB 4; Length 516;
Best Local Similarity 44.4%; Pred. No. 5.6e-27;
Matches 64; Conservative 13; Mismatches 56; Indels 11; Gaps 2;
```

```

QY 1 SNELH-----OVPSNCDLNGGTCVSNKYFSNIHWCNPKKFGGQHCEIDKSKTC 50
Db 23 SGEIHARFRGARSYGCCSEPRFCNGGTCQALYFSDP-VCQCEPGFAGKCCCEIDTRATC 81
QY 51 YEGNGHFYRGKASTDTMGPRCLPWNATVLQOQTYHAHRSDALQLGLGKHNCRPNRNR 110
Db 82 YEDQGISYRGTSWTAESGAECTWNSSALAKPYSGRRPDAIRLGLGNHNYCRNPDRD 141
QY 111 PWCYQVQGLKPLVQECMVHDCADG 134
Db 142 PWCYVFKAGYSSEFCSTPACSEG 165

```

```

RESULT 7
Q91VP2 PRELIMINARY; PRT; 559 AA.
ID Q91VP2
AC Q91VP2;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
GN PLAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC011256; AAH11256.1; -.
DR HSSP; P00761; IAN1.
DR MGD; MGI:97810; Plat.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibronctn1.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01185; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 559 AA; 63122 MW; 8CCE2BDB94514D9 CRC64;

```

```

Query Match 39.8%; Score 315.5; DB 11; Length 559;
Best Local Similarity 44.5%; Pred. No. 3.6e-26;
Matches 61; Conservative 15; Mismatches 56; Indels 5; Gaps 2;
QY 3 ELHQVP----SNCDLNGGTCVSNKYFSNIHWCNPKKFGGQHCEIDKSKTCYEGNGHFY 58
Db 74 QCHSVPRVSRSEPRFCNGGTCQALYFSDP-VCQCPDGFVGKRCIDTRATCFEEQGIT 132

```

```

QY 59 RGKASTDTMGPRCLPWNATVLQOQTYHAHRSDALQLGLGKHNCRPNRNR 118
Db 133 RGTWTAESGAECTWNSSVLKPYNARRPNAIKLGLGNHNYCRNPDRDLKPWCYVFK 192
QY 119 LKPLVQECMVHDCADGK 135
Db 193 GKYTFEFCSTPACPKG 209

```

```

RESULT 8
Q9XT70 PRELIMINARY; PRT; 214 AA.
ID Q9XT70
AC Q9XT70;
DT 01-NOV-1999 (T-EMBLrel. 12, Created)
DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yin J., Idell S.;
RL "Partial mRNA of rabbit uPA.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; AAD39351.1; -.
DR HSSP; P00749; IEGN.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
KW NON TER 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

```

```

Query Match 39.5%; Score 313; DB 6; Length 214;
Best Local Similarity 78.3%; Pred. No. 2.4e-26;
Matches 54; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
QY 67 MGRPCLPWNATVLQOQTYHAHRSDALQLGLGKHNCRPNRNR 126
Db 1 MDRPCLAWNANVLTKTYHAHRPDALQLGLGKHNCRPNRPHQRPWCYVQGLKQLQEC 60
QY 127 MVHDCADGK 135
Db 61 KVHDSGK 69

```

```

RESULT 9
Q8SQ23 PRELIMINARY; PRT; 562 AA.
ID Q8SQ23
AC Q8SQ23;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Enamel organ;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AAM00297.1; -.
DR HSP; P00761; 1ANL.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00051; kringle; 2.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR PROSITE; PS00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PROSITE; PS00136; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR SMART; SM00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 562 AA; 63668 MW; F9B6B4C77CB101E8 CRC64;

Query Match 39.2%; Score 310.5; DB 6; Length 562;
Best Local Similarity 43.8%; Pred. No. 1.3e-25;
Matches 60; Conservative 14; Mismatches 58; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCVGNHGFY 58
Db 77 QCHSVFVKSCSPRCFNGTCTQAIYFSDF-VQCQPVGFGRQCEIDARATCYEDGITY 135
QY 59 RKGASTDTMGRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQV 118
Db 136 RGTWSTESGAEVCNWNNTSGLASMPYNGRRPDAVKLGLGNHNYCRNPKDCKSPWCYIFKA 195
QY 119 LKPLVQECMVHDCADCK 135
Db 196 EKYSDFCSTPACTREK 212

RESULT 10
Q8K0B1 PRELIMINARY; PRT; 564 AA.
ID Q8K0B1
AC Q8K0B1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tissue-type plasminogen activator.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AY029518; AAK40240.1; -.

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Enamel organ;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AAM00297.1; -.
DR HSP; P00761; 1ANL.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00051; kringle; 2.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR PROSITE; PS00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 564 AA; 62726 MW; 459D8BAC6D4A937C CRC64;

Query Match 38.0%; Score 301.5; DB 6; Length 564;
Best Local Similarity 48.7%; Pred. No. 1.2e-24;
Matches 56; Conservative 13; Mismatches 41; Indels 5; Gaps 3;

QY 5 HQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCVGNHGFYRG 60
Db 80 HSPVQSCSEPRCLNGGTCQAIYFSDF-VQCQPEGFVGRCEVDTRARCYEDRGIGYRG 138
QY 61 KASTDTMGRCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYV 115
Db 139 TWSSTESGACVWNNTSGLASMPYNGRRPDAVKLGLGNHNYCRNPKDCKSPWCYV 193

RESULT 11
Q8K0D2 PRELIMINARY; PRT; 517 AA.
ID Q8K0D2
AC Q8K0D2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC031775; AAH31775.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.

```



PRINTS; PRO0018; KRINGLE; 1.  
ProDom; PD000395; Kringle; 1.  
SMART; SM00181; EGF; 3.  
SMART; SM00130; KR; 1.  
SMART; SM00020; TRYPSIN; 1.  
PROSITE; PS00022; EGF 1; 3.  
PROSITE; PS01186; EGF 2; 2.  
PROSITE; PS00021; KRINGLE 1; 1.  
PROSITE; PS00070; KRINGLE 2; 1.  
PROSITE; PS00240; TRYPSIN; 1.  
PROSITE; PS00134; TRYPSIN; 1.  
PROSITE; PS00135; TRYPSIN; 1.  
PROSITE; PS00135; TRYPSIN; 1.  
Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.  
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.  
SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match 33.4%; Score 264.5; DB 11; Length 517;  
Best Local Similarity 42.5%; Pred. No. 1.3e-20;  
Matches 51; Conservative 18; Mismatches 48; Indels 3; Gaps 3;

QY 13 CLNGGTCVSNKYFNSNIHWCNPKFGQHCCEIDKSKTCYEGNGHGYRGKASTDTWGRPCL 72  
DB 116 CQNGGVCSRRRSRFP-TACPDQYKGFCEIGP-DCVGDGYSYRGKASTDTWGRPCL 173  
QY 73 PMSATVLOQTYHAHRSALQGLGKHNYCRNPDRRPPWCYVQVGLKPLVQE-CMVHDC 131  
DB 174 YNNSHLLQETYNMFEDAEATHGIEHNFERNPDGDHFKPFCVKNSEKVKWEYCDTVTC 233

RESULT 12  
Q14520 PRELIMINARY; PRT; 560 AA.  
ID Q14520  
AC Q14520; O00663;  
DT 01-NOV-1996 (T-EMBLrel. 01, Created)  
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
DE HGF activator like protein (Hyaluronan binding protein 2).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kitamura N.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=96425001; PubMed=8827452;  
RX Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T., Tomita M.;  
RA "Purification and characterization of a novel hyaluronan-binding protein (PHBP) from human plasma: it has three EGF, a kringle and a serine protease domain, similar to hepatocyte growth factor activator.";  
RL J. Biochem. 119:1157-1165(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon, and Kidney;  
RA Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; D43742; BAA08576.1; -;  
DR EMBL; BC031412; AAB46909.1; -;  
DR HSSP; P00763; 1DPO.  
DR MEROPS; S01.033; -;  
DR Genew; HGNC-4798; HAP2.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00008; EGF; 3.

PRINTS; PRO0018; KRINGLE; 1.  
ProDom; PD000395; Kringle; 1.  
SMART; SM00181; EGF; 3.  
SMART; SM00130; KR; 1.  
SMART; SM00020; TRYPSIN; 1.  
PROSITE; PS00022; EGF 1; 3.  
PROSITE; PS01186; EGF 2; 2.  
PROSITE; PS00021; KRINGLE 1; 1.  
PROSITE; PS00070; KRINGLE 2; 1.  
PROSITE; PS00240; TRYPSIN; 1.  
PROSITE; PS00134; TRYPSIN; 1.  
PROSITE; PS00135; TRYPSIN; 1.  
PROSITE; PS00135; TRYPSIN; 1.  
Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.  
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.  
SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match 33.4%; Score 264.5; DB 11; Length 517;  
Best Local Similarity 42.5%; Pred. No. 1.3e-20;  
Matches 51; Conservative 18; Mismatches 48; Indels 3; Gaps 3;

QY 13 CLNGGTCVSNKYFNSNIHWCNPKFGQHCCEIDKSKTCYEGNGHGYRGKASTDTWGRPCL 72  
DB 116 CQNGGVCSRRRSRFP-TACPDQYKGFCEIGP-DCVGDGYSYRGKASTDTWGRPCL 173  
QY 73 PMSATVLOQTYHAHRSALQGLGKHNYCRNPDRRPPWCYVQVGLKPLVQE-CMVHDC 131  
DB 174 YNNSHLLQETYNMFEDAEATHGIEHNFERNPDGDHFKPFCVKNSEKVKWEYCDTVTC 233

RESULT 12  
Q14520 PRELIMINARY; PRT; 560 AA.  
ID Q14520  
AC Q14520; O00663;  
DT 01-NOV-1996 (T-EMBLrel. 01, Created)  
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
DE HGF activator like protein (Hyaluronan binding protein 2).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kitamura N.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=96425001; PubMed=8827452;  
RX Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T., Tomita M.;  
RA "Purification and characterization of a novel hyaluronan-binding protein (PHBP) from human plasma: it has three EGF, a kringle and a serine protease domain, similar to hepatocyte growth factor activator.";  
RL J. Biochem. 119:1157-1165(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon, and Kidney;  
RA Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; D43742; BAA08576.1; -;  
DR EMBL; BC031412; AAB46909.1; -;  
DR HSSP; P00763; 1DPO.  
DR MEROPS; S01.033; -;  
DR Genew; HGNC-4798; HAP2.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00008; EGF; 3.

PRINTS; PRO0018; KRINGLE; 1.  
ProDom; PD000395; Kringle; 1.  
SMART; SM00181; EGF; 3.  
SMART; SM00130; KR; 1.  
SMART; SM00020; TRYPSIN; 1.  
PROSITE; PS00022; EGF 1; 3.  
PROSITE; PS01186; EGF 2; 2.  
PROSITE; PS00021; KRINGLE 1; 1.  
PROSITE; PS00070; KRINGLE 2; 1.  
PROSITE; PS00240; TRYPSIN; 1.  
PROSITE; PS00134; TRYPSIN; 1.  
PROSITE; PS00135; TRYPSIN; 1.  
PROSITE; PS00135; TRYPSIN; 1.  
Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.  
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.  
SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match 33.4%; Score 264.5; DB 11; Length 517;  
Best Local Similarity 42.5%; Pred. No. 1.3e-20;  
Matches 51; Conservative 18; Mismatches 48; Indels 3; Gaps 3;

QY 13 CLNGGTCVSNKYFNSNIHWCNPKFGQHCCEIDKSKTCYEGNGHGYRGKASTDTWGRPCL 72  
DB 116 CQNGGVCSRRRSRFP-TACPDQYKGFCEIGP-DCVGDGYSYRGKASTDTWGRPCL 173  
QY 73 PMSATVLOQTYHAHRSALQGLGKHNYCRNPDRRPPWCYVQVGLKPLVQE-CMVHDC 131  
DB 174 YNNSHLLQETYNMFEDAEATHGIEHNFERNPDGDHFKPFCVKNSEKVKWEYCDTVTC 233

RESULT 12  
Q14520 PRELIMINARY; PRT; 560 AA.  
ID Q14520  
AC Q14520; O00663;  
DT 01-NOV-1996 (T-EMBLrel. 01, Created)  
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
DE HGF activator like protein (Hyaluronan binding protein 2).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kitamura N.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=96425001; PubMed=8827452;  
RX Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T., Tomita M.;  
RA "Purification and characterization of a novel hyaluronan-binding protein (PHBP) from human plasma: it has three EGF, a kringle and a serine protease domain, similar to hepatocyte growth factor activator.";  
RL J. Biochem. 119:1157-1165(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon, and Kidney;  
RA Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; D43742; BAA08576.1; -;  
DR EMBL; BC031412; AAB46909.1; -;  
DR HSSP; P00763; 1DPO.  
DR MEROPS; S01.033; -;  
DR Genew; HGNC-4798; HAP2.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00008; EGF; 3.

PRINTS; PRO0018; KRINGLE; 1.  
ProDom;

```
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase;
KW Serine protease.
SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBDIA9 CRC64;

Query Match 32.4%; Score 257; DB 6; Length 616;
Best Local Similarity 40.9%; Pred.No.1e-19;
Matches 52; Conservative 14; Mismatches 53; Indels 8; Gaps 3;

QY 6 QVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFGYRGKASTD 65
DB 176 QVCSINPLNGSGCLQTE---GRLRCRPTGYAGRLCDVDLKERCYSDRGLSYRGMAQTT 232

QY 66 TWGRPCLPNNSATVLQOYTHAHRSDALQGLGKHNYCRNPDRRRPWCYVQVGLKPLVQ 124
DB 233 LSGAPQCPWAS----BATYWNMTAEQALNWGLGDHAFRCRPNDRPTREPCFVWRGDQLSWQ 288

QY 125 EGMVHDC 131
DB 289 YCLARC 295

RESULT 14
Q8VCS4 PRELIMINARY; PRT; 653 AA.
AC Q8VCS4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 70.6 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -.
DR HSSP; P00761; 1AN1.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibnctnl.
DR InterPro; IPR0000562; FN Type II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR000722; CHYMOTRYPsin.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 31.6%; Score 250.5; DB 11; Length 653;
Best Local Similarity 42.0%; Pred.No.5.6e-19;
Matches 50; Conservative 11; Mismatches 41; Indels 17; Gaps 2;

QY 5 HQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGH 56
DB 239 HTACLSPCLNGGTC-----HLIVGTGTSVCTCPLGYAGRCFNIVPTEHCFLNGT 289

QY 57 FYRGKASTDTMGRCPLPWNNSATVLQOYTHAHRSDALQGLGKHNYCRNPDRRRPWCYV 115
DB 290 EYRGVASTAASGLSCLAMNSDLLYQELHVDVAAAVLLGLGPHAYCRNPDKDRPWCYV 348

RESULT 15
O35727 PRELIMINARY; PRT; 597 AA.
AC O35727;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Factor XII.
GN F12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Schloesser M.; Schwager S.; Engel W.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; X99571; CAA67891.1; -.
DR HSSP; P00760; 1AQ7.
DR MEROPS; S01.211; -.
DR MGD; MGI:1891012; F12.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibnctnl.
DR InterPro; IPR0000562; FN Type II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR000722; CHYMOTRYPsin.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
```

```

DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;

Query Match 31.4%; Score 249; DB 11; Length 597;
Best Local Similarity 40.8%; Pred. No. 7.4e-19;
Matches 49; Conservative 14; Mismatches 49; Indels 8; Gaps 3;

QY 13 CLNGGTCVSNKPSNIHWNCPCPKFGQHCHEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
183 CLNGGSC---LVEDHPLCRCPYGYGFCDLDMATCYEGRGLSYRGAGTTQSGAPCQ 239

QY 73 PMSATVLOQTY-HAHRSDALQLGLGKHNYCRNPNRRPWCYVQVLKPLVQECMVHDC 131
240 RW----TVEATYRNMTKEQALSGLGHAFRCRPNDDTRPWCYVSGDRLSWDYCGLEQC 295

RESULT 16
Q81Z25 PRELIMINARY; PRT; 615 AA.
AC Q81Z25;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Coagulation factor XII-Mie.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SQ SEQUENCE FROM N.A.

Query Match 29.4%; Score 233; DB 4; Length 615;
Best Local Similarity 42.3%; Pred. No. 4.3e-17;
Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;

QY 13 CLNGGTCVSNKPSNIHWNCPCPKFGQHCHEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
183 CLNGGRCLE---VEGHLRCHCPVGYTFPCDVTKASCYDGRGLSYRGLARTILSGAPCQ 239

QY 73 PMSATVLOQTY-HAHRSDALQLGLGKHNYCRNPNRRPWCYV 115
240 PWAS---EATYRNVTAEQARNWGLGGHAFRCRPNDDIRPWCYV 279

RESULT 17
Q9BZW1 PRELIMINARY; PRT; 395 AA.
AC Q9BZW1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neonatal thrombolytic agent alpha-form (Fragment).
OS Homo sapiens (Human).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF260825; AAK11956.1; -.
DR HSSP; P00750; 1PK2
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00019; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 28.9%; Score 229.5; DB 4; Length 395;
Best Local Similarity 42.7%; Pred. No. 6.5e-17;
Matches 44; Conservative 10; Mismatches 44; Indels 5; Gaps 2;

QY 30 WNCPCPKFGQHCHEIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWSATVLOQTYHAHRS 89
33 WCNCS---GRAQCS-EGNSDCYFGNGSAVRGTHSLTESGASCLPWSMILIGKVTYQAQP 87

QY 90 DALQLGLGKHNYCRNPNRRPWCYVQVLKPLVQECMVHDC 132
88 SAQALGLGKHNYCRNPDGDAKPMCHVLKNRRLTWECYCDVPSCS 130

RESULT 18
Q8NG20 PRELIMINARY; PRT; 90 AA.
AC Q8NG20;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen/activator kringle.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "Production of kringle fragment.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF282882; AAM52248.1; -.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.

```

```
DR PROS:TE: PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
SQ SEQUENCE 90 AA; 9804 MW; A33887F9DF4C7B1 CRC64;

Query Match
Best Local Similarity 27.6%; Score 218.5; DB 4; Length 90;
Matches 41; Conservative 5; Mismatches 36; Indels 1; Gaps 1;

QY 50 CYEGNGHFRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDNRR 109
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
8 CYFGNGSAYRGTHSLTESGASCLPWN SMILIGKVYTAQNP SAQALGLGKHNYCRNPDGDA 67
QY 110 RPWCYVQVGLKPLVQECMVHDC 132
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
68 KPWCYT-TNPKLYDYCDVPPQCA 89

RESULT 19
Q00675 PRELIMINARY; PRT; 202 AA.
AC Q00675
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tissue-type plasminogen activator (Fragment).
GN TPA.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RX MEDLINE=97199025; PubMed=9047000;
RA Johnson A.L., Bridgham J.T., Anthony R.V.;
RT "Expression of avian urokinase and tissue-type plasminogen activator
RT messenger ribonucleic acid during follicle development and atresia.";
RL Biol. Reprod. 56:581-588(1997).
CC !- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; U31988; AAN74955.1; -.
DR MEROPS; S01.232; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00440; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;

Query Match
Best Local Similarity 24.9%; Score 197.5; DB 13; Length 202;
Matches 41; Conservative 11; Mismatches 50; Indels 9; Gaps 2;

QY 30 WCNCPK--KFGGHC-----IDSKTCYEGNGHFRGKASTDTMGRCPLPWN SATVL 80
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
12 WCYVFAGKGI SEFCSTPACTKVAEDGDCYTGNGLA YRGTRSRTSKGSFCLPWN PVFLT 71
QY 81 QOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLKPLVQECMVHDC 131
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
72 SKIYTALEQRALGLGKHNCNCPDGDAPQWCHWKDRQLTW EYCDVQC 122

RESULT 20
```

```
Q8AXX3 PRELIMINARY; PRT; 421 AA.
AC Q8AXX3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Kremen2.
GN KRM2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]_TaxID=8355;
RP SEQUENCE FROM N.A.
RX MEDLINE=22308873; PubMed=12421700;
RA Davidson G., Mao B., Del Barco Barrantes I., Niehrs C.;
RT "Kremen proteins interact with Dickkopf1 to regulate anteroposterior
RT CNS patterning.";
RL Development 129:5587-5596(2002).
DR EMBL; AY150813; AAN64661.1; -.
SQ SEQUENCE 421 AA; 46274 MW; 74D04DB1682CD15F CRC64;

Query Match
Best Local Similarity 21.9%; Score 174; DB 13; Length 421;
Matches 36; Conservative 6; Mismatches 28; Indels 6; Gaps 3;

QY 41 HCEIDKSKTCYEGNGHFRGKAS-TDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGH 99
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23 HPFLSE---CFTVNGRDYRGTVSQA GPGTGPLYNNQTT--QHL YNAQSDPDGELGLGNH 77
QY 100 NYCRRNPDNRRPWCYV 115
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
78 NYCRRNPDADVPWCYV 93

RESULT 21
Q8AXY6 PRELIMINARY; PRT; 947 AA.
AC Q8AXY6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Muscle-specific receptor tyrosine kinase MusK.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RX MEDLINE=20538710; PubMed=11083926;
RA Ip F.C., Glass D.G., Gies D.R., Cheung J., Lai K.O., Fu A.K.;
RA Yancopoulos G.D., Ip N.Y.;
RT "Cloning and characterization of muscle-specific kinase in chicken.";
RL Mol. Cell. Neurosci. 16:661-673(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Gies D., Glass D.J., Yancopoulos G.D.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY143173; AAN05008.1; -.
KW Receptor; Kinase.
SQ SEQUENCE 947 AA; 105588 MW; E6C3FCC5796CC8BE CRC64;

Query Match
Best Local Similarity 21.5%; Score 170.5; DB 13; Length 947;
Matches 46; Conservative 13; Mismatches 34; Indels 49; Gaps 8;

QY 4 LHQVFNCDCLNGGTCVSNKYFSNHWNCMPKFGGQHC EIDKSKTCYEGNGHFRGKAS 63
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
441 LHQDFSA-----THIFFDFKKE-----NITRTC YSGNGQFYQGWAN 478

QY 64 TDTMGRCPLPWN SATVLQOQTYHAHRS-----SDALQLGLGKHNYCRNPDNRRPWC 113
```

```
Db 479 VTASGIPCKWS-----DOAPHLHRRTPQVPFLSDA-----ENYCRNPGGENRPMWC 526
QY 114 YVQVGLKPLV--QECMVHDCAD 133
Db 527 YTK---DPSVTWEYCSVSPCGD 545

RESULT 22
Q91691
ID Q91691 PRELIMINARY; PRT; 716 AA.
AC Q91691;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Growth factor Livertine.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Ruiz i Altaba A., Thery C.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR HSSP; P00747; 1CEA.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 716 AA; 81971 MW; 508376A0E4398798 CRC64;

Query Match 20.9%; Score 166; DB 13; Length 716;
Best Local Similarity 31.1%; Pred. No. 1.1e-09;
Matches 42; Conservative 11; Mismatches 40; Indels 42; Gaps 6;

QY 30 WC-----NCPKFGQGHCEIDKSK-----TCYEGNGHFYRGKASTDTMGRCLP 73
Db 258 WCYTDDNVEKEP-----CRITCKQRLSNIETSTCFKRGEGYRKANTTISGIPQOR 313
QY 74 WNSAT-----VLQOTYHAHRSALQLGLGKHNYCRNPDNRNRRPWCYVQV-----GL 119
Db 314 WDSQTPQSHRFLPEKYPCKGLD-----ENYCRNPDGSEAPWCFITLFCMRMAYCFQI 365
QY 120 KPLVQECMVHDCADG 134
Db 366 KRCKDDVLEPDCVHG 380

RESULT 23
Q8WMR1
ID Q8WMR1 PRELIMINARY; PRT; 359 AA.
AC Q8WMR1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
```

```
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Pirie-Shepherd S.R., Coffman K.T., Resnick D., Chan R., Kisker O.,
RA Folkman J., Waters D.J.;
RT "Angiostatin is Detectable in the Urine of Dogs with Spontaneous Bone
RT Cancer.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; AY069985; AAL58519.1; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003966; Prothrombin.
DR Pfam; PF00051; kringle; 4.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
KW Glycoprotein; Kringle.
FT NON_TER 1
FT NON_TER 359
SQ SEQUENCE 359 AA; 41172 MW; 776D35F4AB0BDD9E CRC64;

Query Match 20.9%; Score 165.5; DB 6; Length 359;
Best Local Similarity 29.0%; Pred. No. 6.1e-10;
Matches 42; Conservative 13; Mismatches 43; Indels 47; Gaps 7;

QY 5 HQVPSNCDCLNGTCTVSNKYFSN-----IHWNC-NCPKFGQGHCEIDKSKT-----49
Db 211 NRTFENFPCKN-----LDENYCRNPDGETAPWCYTTSNEVEHHCQIPSCSSPITTEYLD 266
QY 50 -----CYEGNGHFYRGKASTDTMGRCLPWN SATVLQOTYHAHRSAL--92
Db 267 APASVPPEQTPVQECYHGNGQSYRGTSSTITGRKQCSWSMT-----PHRHEKTP 319
QY 93 ---QLGLGKHNYCRNPDNRNRRPWCY 114
Db 320 HPFEAGL-TWNYCRNPDADKSPWCY 343

RESULT 24
Q9ROW3
ID Q9ROW3 PRELIMINARY; PRT; 812 AA.
AC Q9ROW3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bangert K., Johnsen A.H., Thorsen S.;
RT "Rat plasminogen: cDNA and gene structure.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91250378; PubMed=1645711;
RA Kanakas J.J., Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).
```



```

Db 236 DETSSVFKPRDLRCDECEILENLVLCQTEYIFARSNPMLRLKLPNCEDLQPESPEAAN 295
Qy 42 C-----EIDKSKTCYEGNGHYRGKASTDTMGPRCLPWNATVLIQYTHAHRSDAL 92
Db 296 CIRIGIPMADPINKNKKYNSGTVDYRGTVSVTKSGROCPWNS-----QYPHTHSFTAL 350
Qy 93 QLG--LGKHNYCRNPDNR--PWCY 114
Db 351 RFPELNGGHSYCRNPGNKEAPWCF 375

RESULT 27
Q8BNP9
ID Q8BNP9 PRELIMINARY; PRT; 937 AA.
AC Q8BNP9;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DE Receptor tyrosine kinase-like orphan receptor 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adipose tissue;
RA MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK080896; BAC38068.1; -.
SQ SEQUENCE 937 AA; 104118 MW; CB440323CA66EC28 CRC64;

Query Match 20.0%; Score 158.5; DB 11; Length 937;
Best Local Similarity 29.7%; Pred. No. 1e-08;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

Qy 2 NELHQP-----SNCDCLNGGTGCVSNKYFS-----NIHWCNC-----PKKFGGQH 41
Db 236 DETSSVFKPRDLRCDECEILENLVLCQTEYIFARSNPMLRLKLPNCEDLQPESPEAAN 295
Qy 42 C-----EIDKSKTCYEGNGHYRGKASTDTMGPRCLPWNATVLIQYTHAHRSDAL 92
Db 296 CIRIGIPMADPINKNKKYNSGTVDYRGTVSVTKSGROCPWNS-----QYPHTHSFTAL 350
Qy 93 QLG--LGKHNYCRNPDNR--PWCY 114
Db 351 RFPELNGGHSYCRNPGNKEAPWCF 375

RESULT 28
Q8BG10
ID Q8BG10 PRELIMINARY; PRT; 937 AA.
AC Q8BG10;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DE Receptor tyrosine kinase-like orphan receptor 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adipose tissue;
RA MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
```

```

RL Nature 420:563-573(2002).
DR EMBL; AK046699; BAC32840.1; -.
DR EMBL; AK049369; BAC33714.1; -.
SQ SEQUENCE 937 AA; 104088 MW; D6F2D84E67D03C69 CRC64;

Query Match 20.0%; Score 158.5; DB 11; Length 937;
Best Local Similarity 29.7%; Pred. No. 1e-08;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

Qy 2 NELHQP-----SNCDCLNGGTGCVSNKYFS-----NIHWCNC-----PKKFGGQH 41
Db 236 DETSSVFKPRDLRCDECEILENLVLCQTEYIFARSNPMLRLKLPNCEDLQPESPEAAN 295
Qy 42 C-----EIDKSKTCYEGNGHYRGKASTDTMGPRCLPWNATVLIQYTHAHRSDAL 92
Db 296 CIRIGIPMADPINKNKKYNSGTVDYRGTVSVTKSGROCPWNS-----QYPHTHSFTAL 350
Qy 93 QLG--LGKHNYCRNPDNR--PWCY 114
Db 351 RFPELNGGHSYCRNPGNKEAPWCF 375

RESULT 29
Q46506
ID Q46506 PRELIMINARY; PRT; 454 AA.
AC Q46506;
DT 01-JUN-1998 (TremBLrel. 06, Created)
DT 01-JUN-1998 (TremBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Apolipoprotein a (Fragment).
GN BAAPOA.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF029691; AAB97886.1; -.
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.999; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
KW Serine protease.
FT NON TER 1
SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;

Query Match 19.9%; Score 158; DB 6; Length 454;
Best Local Similarity 39.8%; Pred. No. 5.2e-09;
Matches 35; Conservative 10; Mismatches 31; Indels 12; Gaps 4;

Qy 50 CYEGNGHYRGKASTDTMGPRCLPWNATVLIQ--QYTHAHRSDALQLGLGKHNYCRNPDN 107
```

Db 102 CYHGQCYVGRGFSFTVTGRTCSWSSMTPHQHKTPTENHPNDGLTM-----NYCRNPD 156  
QY 108 RRRPWCYVQVGLKPLV--QECMVHDCAD 133  
Db 157 DTGPWCFT--MDPSVRWEYCNLTRCSD 181

## RESULT 30

Q9UIRS  
ID Q9UIRS PRELIMINARY; PRT; 113 AA.  
AC Q9UIRS;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Apolipoprotein(a) (Fragment).  
GN APOA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21181705; PubMed=11285247;  
RA Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.;  
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV  
RT types 6 to 10 domain affect lip(a) plasma concentrations and have  
RT different patterns in Africans and Caucasians.";  
RL Hum. Mol. Genet. 10:815-824(2001).  
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; AF158663; AAF03680.1; -;  
DR EMBL; AF158662; AAF03680.1; JOINED.  
DR HSSP; P00747; IPMK.  
DR InterPro; IPR000001; Kringle.  
DR Pfam; PF00051; kringle; 1.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
KW Glycoprotein; Kringle; Lipoprotein.  
FT NON\_TER 1 1  
FT NON\_TER 113 113  
SQ SEQUENCE 113 AA; 12685 MW; F3D65681D9B5253A CRC64;

Query Match 19.7%; Score 156; DB 4; Length 113;  
Best Local Similarity 37.8%; Pred. No. 1.9e-09;  
Matches 34; Conservative 12; Mismatches 32; Indels 12; Gaps 4;

QY 48 KTCYEGNGHFYRGKASDTMTGRPCLPWNSATV--LQCTYHAHRS DALQLGLGKHNYCRNP 105  
Db 9 RQCYHNGNSYRGTFTTGTTCQSWSSWTPRHQRTPENYPNDGLTM-----NYCRNP 63

QY 106 DNRREWCYVQVGLKPLV--QECMVHDCAD 133  
Db 64 DADTGPWCFT--MDPSIRWEYCNLTRCSD 90

Search completed: December 3, 2003, 14:42:57  
Job time : 16.6105 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:43:04 ; Search time 13.8597 Seconds  
(without alignments)  
1811.566 Million cell updates/sec

Title: US-09-880-503-4  
Perfect score: 793  
Sequence: 1 SNELHQVPSNCDCLNGGTCV.....QVGLKPLVQCMVHDCADGK 135

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	793	100.0	135	10	US-09-880-503-4
2	793	100.0	138	10	US-09-884-186-12
3	793	100.0	138	15	US-10-237-667-12
4	793	100.0	138	15	US-10-237-708-12
5	793	100.0	138	15	US-10-237-866-12
6	793	100.0	138	15	US-10-237-871-12
7	793	100.0	138	15	US-10-237-624-12
8	793	100.0	143	10	US-09-880-503-8
9	793	100.0	337	15	US-10-106-698-6266
10	793	100.0	403	10	US-09-880-503-6
11	793	100.0	411	10	US-09-880-503-3
12	793	100.0	431	10	US-09-264-468B-1
13	793	100.0	431	12	US-10-301-822-161
14	793	100.0	431	12	US-10-131-985-21
15	793	100.0	431	14	US-10-076-421-2

16	793	100.0	431	15	US-10-171-311-184	Sequence 194, Appl
17	793	100.0	431	15	US-10-193-656-4	Sequence 4, Appl
18	793	99.7	431	12	US-10-247-671-149	Sequence 149, Appl
19	783	98.7	411	12	US-10-407-821-2	Sequence 2, Appl
20	510	64.3	88	10	US-09-880-503-1	Sequence 1, Appl
21	510	64.3	96	10	US-09-880-503-9	Sequence 9, Appl
22	328.5	41.4	527	11	US-09-987-457-18	Sequence 18, Appl
23	328.5	41.4	527	11	US-09-987-455-19	Sequence 19, Appl
24	328.5	41.4	562	9	US-09-969-271-7	Sequence 7, Appl
25	328.5	41.4	562	10	US-09-974-298-145	Sequence 145, Appl
26	328.5	41.4	562	12	US-10-443-701-4	Sequence 4, Appl
27	328.5	41.4	562	15	US-10-193-656-8	Sequence 8, Appl
28	293	36.9	49	12	US-10-349-543-5	Sequence 5, Appl
29	288	36.3	323	10	US-09-880-503-7	Sequence 7, Appl
30	283	35.7	47	10	US-09-880-503-2	Sequence 2, Appl
31	268	33.8	44	12	US-10-349-543-1	Sequence 1, Appl
32	260.5	32.8	655	15	US-10-172-712-28	Sequence 28, Appl
33	257.5	32.5	560	10	US-09-912-559-3	Sequence 3, Appl
34	257.5	32.5	560	10	US-09-912-559-4	Sequence 4, Appl
35	257.5	32.5	560	15	US-10-172-712-32	Sequence 32, Appl
36	233	29.4	615	11	US-09-858-909-2	Sequence 2, Appl
37	233	29.4	615	15	US-10-172-712-30	Sequence 30, Appl
38	220.5	27.8	326	14	US-10-057-951-3	Sequence 3, Appl
39	220.5	27.8	372	9	US-09-084-491A-3	Sequence 3, Appl
40	220.5	27.8	372	14	US-10-102-704-3	Sequence 3, Appl
41	215	27.1	354	11	US-09-987-457-10	Sequence 10, Appl
42	215	27.1	354	11	US-09-987-455-11	Sequence 11, Appl
43	215	27.1	377	11	US-09-987-455-8	Sequence 8, Appl
44	176	22.2	343	11	US-09-987-457-14	Sequence 14, Appl
45	176	22.2	343	11	US-09-987-455-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1  
US-09-880-503-4  
; Sequence 4, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; TITLE OF INVENTION: TISSUE CONTRACTILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 135  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-4

Query Match		100.0%	Score 793;	DB 10;	Length 135;
Best Local Similarity		100.0%	Pred. No. 2.6e-72;		
Matches 135;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG	60		
Db	1	SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG	60		
Qy	61	KASDTMTGRCLPWSATVILQOTYHAHRSALQGLGKHVCRPNDRRRPWCVVQVGLK	120		
Db	61	KASDTMTGRCLPWSATVILQOTYHAHRSALQGLGKHVCRPNDRRRPWCVVQVGLK	120		
Qy	121	PLVQECMVHDCADGK	135		
Db	121	PLVQECMVHDCADGK	135		

## RESULT 2

US-09-984-186-12  
; Sequence 12, Application US/09984186  
; Patent No. US20020151011A1  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; Fournier, Alain  
; Guittion, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (Patentin)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/984,186  
FILING DATE: 29-Oct-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 138 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-984-186-12

Query Match 100.0%; Score 793; DB 10; Length 138;  
Best Local Similarity 100.0%; Pred. No. 2.6e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60

Db 4 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 63

Qy 61 KASDTDMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120

Db 64 KASDTDMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 123

Qy 121 PLVQECMVHDCADGK 135

Db 124 PLVQECMVHDCADGK 138

## RESULT 4

## RESULT 3

US-10-237-667-12  
; Sequence 12, Application US/10237667  
; Publication No. US20030022308A1  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; Fournier, Alain  
; Guittion, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (Patentin)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,667  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 138 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-237-667-12

Query Match 100.0%; Score 793; DB 15; Length 138;  
Best Local Similarity 100.0%; Pred. No. 2.6e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60

Db 4 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 63

Qy 61 KASDTDMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120

Db 64 KASDTDMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 123

Qy 121 PLVQECMVHDCADGK 135

Db 124 PLVQECMVHDCADGK 138

```
US-10-237-708-12
; Sequence 12, Application US/10237708
; Publication No. US20030036170A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
;             Fournier, Alain
;             Guitton, Jean-Dominique
;             Jung, Gerard
;             Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
;
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,708
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12

Query Match      100.0%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.6e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 63

Qy 61 KASDTDTGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 64 KASDTDTGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 123

Qy 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 5
US-10-237-866-12
; Sequence 12, Application US/10237866
; Publication No. US20030036171A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
;             Fournier, Alain
;             Guitton, Jean-Dominique
;             Jung, Gerard
;             Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
;
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,866
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match      100.0%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.6e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 63

Qy 61 KASDTDTGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 64 KASDTDTGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 123

Qy 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 6
US-10-237-871-12
; Sequence 12, Application US/10237871
```

```
; Publication No. US20030036172A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
;
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,871
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
;
; US-10-237-871-12
;
; Query Match 100.0%; Score 793; DB 15; Length 138;
; Best Local Similarity 100.0%; Pred. No. 2.6e-72;
; Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
; Db 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 63
;
; Qy 61 KASDTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
; Db 64 KASDTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVQVGLK 123
;
; Qy 121 PLVQECMVHDCADGK 135
; Db 124 PLVQECMVHDCADGK 138
;
; RESULT 7
; US-10-237-871-12
; Sequence 12, Application US/10237624
; Publication No. US20030082747A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
;
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,624
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
;
; US-10-237-624-12
;
; Query Match 100.0%; Score 793; DB 15; Length 138;
; Best Local Similarity 100.0%; Pred. No. 2.6e-72;
; Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
; Db 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 63
;
; Qy 61 KASDTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
; Db 64 KASDTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVQVGLK 123
;
; Qy 121 PLVQECMVHDCADGK 135
; Db 124 PLVQECMVHDCADGK 138
;
; RESULT 8
; US-09-880-503-8
; Sequence 8, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
```

; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; TITLE OF INVENTION: TISSUE CONTRACTABILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880, 503  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 143  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-8

Query Match 100.0%; Score 793; DB 10; Length 143;  
Best Local Similarity 100.0%; Pred. No. 2.7e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFGYRG 60  
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFGYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 120  
Db 61 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
Db 121 PLVQECMVHDCADGK 135

## RESULT 9

US-10-106-698-6266  
; Sequence 6266, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; PRIOR FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 6266  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-106-698-6266

Query Match 100.0%; Score 793; DB 15; Length 337;  
Best Local Similarity 100.0%; Pred. No. 6.9e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFGYRG 60  
Db 27 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFGYRG 86  
QY 61 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 120  
Db 87 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 146  
QY 121 PLVQECMVHDCADGK 135  
Db 147 PLVQECMVHDCADGK 161

## RESULT 10

US-09-880-503-6  
; Sequence 6, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; TITLE OF INVENTION: TISSUE CONTRACTABILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880, 503  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 403  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-6

Query Match 100.0%; Score 793; DB 10; Length 403;  
Best Local Similarity 100.0%; Pred. No. 8.3e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFGYRG 60  
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFGYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 120  
Db 61 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
Db 121 PLVQECMVHDCADGK 135

## RESULT 11

US-09-880-503-3  
; Sequence 3, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; TITLE OF INVENTION: TISSUE CONTRACTABILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880, 503  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-3

Query Match 100.0%; Score 793; DB 10; Length 411;  
Best Local Similarity 100.0%; Pred. No. 8.5e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFGYRG 60  
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFGYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 120

Db 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVYQVGLK 120

QY 121 PLVQECMVHDCADGK 135

Db 121 PLVQECMVHDCADGK 135

## RESULT 12

US-09-264-468B-1

Sequence 1, Application US/09264468B

Patent No. US20020106775A1

GENERAL INFORMATION:

APPLICANT: Wang, Jieyi

APPLICANT: Nienaber, Vicki L.

APPLICANT: Henkin, Jack

APPLICANT: Smith, Richard A.

APPLICANT: Walter, Karl A.

APPLICANT: Severin, Jean M.

APPLICANT: Edalji, Rohinton

APPLICANT: Johnson Jr., Robert W.

APPLICANT: Holzman, Thomas F.

TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE

FILE REFERENCE: 6310.US.P1

CURRENT APPLICATION NUMBER: US/09/264,468B

PRIOR FILING DATE: 1998-03-05

PRIOR FILING DATE: 1998-03-06

NUMBER OF SEQ ID NOS: 23

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 431

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SIGNAL

LOCATION: (1)...(20)

OTHER INFORMATION: Leader sequence

NAME/KEY: VARIANT

LOCATION: (279)...(279)

OTHER INFORMATION: Xaa = any amino acid

NAME/KEY: VARIANT

LOCATION: (302)...(302)

OTHER INFORMATION: Xaa = any amino acid

US-09-264-468B-1

Query Match 100.0%; Score 793; DB 10; Length 431;

Best Local Similarity 100.0%; Pred. No. 9e-72;

Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60

Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVYQVGLK 120

Db 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVYQVGLK 140

QY 121 PLVQECMVHDCADGK 135

Db 141 PLVQECMVHDCADGK 155

## RESULT 13

US-10-301-822-161

Sequence 161, Application US/10301822

Publication No. US20030148410A1

GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc.

APPLICANT: Berger, Allison

APPLICANT: Guillemette, Tracy L.

APPLICANT: Kamatkar, Shubhangi

APPLICANT: Schlegel, Robert

APPLICANT: Monahan, John E.

Query Match 100.0%; Score 793; DB 12; Length 431;

Best Local Similarity 100.0%; Pred. No. 9e-72;

Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60

Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNKCRNPDRRPPWCYVQVGLK 120  
Db 81 KASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNKCRNPDRRPPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155

RESULT 15  
US-10-076-421-2  
; Sequence 2, Application US/10076421  
; Publication No. US20020193304A1  
; GENERAL INFORMATION:  
; APPLICANT: WADA, MANABU  
; APPLICANT: WADA, NAKAO  
; TITLE OF INVENTION: ANTI-HIV AGENTS  
; FILE REFERENCE: HAYAK-9  
; CURRENT APPLICATION NUMBER: US/10/076,421  
; CURRENT FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: JP 2001-42655  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: JP 2001-184284  
; PRIOR FILING DATE: 2001-06-19  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-076-421-2

Query Match 100.0%; Score 793; DB 14; Length 431;  
Best Local Similarity 100.0%; Pred. No. 9e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHOVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGOHCEIDKSKTCYEGNGHFRG 60  
Db 21 SNELHOVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGOHCEIDKSKTCYEGNGHFRG 80

QY 61 KASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNKCRNPDRRPPWCYVQVGLK 120  
Db 81 KASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNKCRNPDRRPPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155

RESULT 16  
US-10-171-311-184  
; Sequence 184, Application US/10171311  
; Publication No. US20030087270A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Chen, Yan  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Monahan, John  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Glatt, Karen  
; APPLICANT: Gannavarapu, Manjula  
; APPLICANT: Hoerish, Sebastian  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY  
; FILE REFERENCE: MRI-035  
; CURRENT APPLICATION NUMBER: US/10/171,311  
; CURRENT FILING DATE: 2002-06-12  
; PRIOR APPLICATION NUMBER: US 60/298,159  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,155  
; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/335,936  
; PRIOR FILING DATE: 2001-11-14  
; NUMBER OF SEQ ID NOS: 238  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 184  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-171-311-184

Query Match 100.0%; Score 793; DB 15; Length 431;  
Best Local Similarity 100.0%; Pred. No. 9e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHOVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGOHCEIDKSKTCYEGNGHFRG 60  
Db 21 SNELHOVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGOHCEIDKSKTCYEGNGHFRG 80

QY 61 KASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNKCRNPDRRPPWCYVQVGLK 120  
Db 81 KASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNKCRNPDRRPPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155

RESULT 17  
US-10-193-656-4  
; Sequence 4, Application US/10193656  
; Publication No. US20030096733A1  
; GENERAL INFORMATION:  
; APPLICANT: NY, Tor  
; APPLICANT: HOLMDAHL, Rikard  
; APPLICANT: LI, Jinan  
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS  
; FILE REFERENCE: 3810/LJ577-US3  
; CURRENT APPLICATION NUMBER: US/10/193,656  
; CURRENT FILING DATE: 2002-07-10  
; PRIOR APPLICATION NUMBER: US 60/304,461  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/304,490  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/305,182  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: GenBank / P00749  
; DATABASE ENTRY DATE: 1986-07-21  
; RELEVANT RESIDUES: (1)..(431)  
US-10-193-656-4

Query Match 100.0%; Score 793; DB 15; Length 431;  
Best Local Similarity 100.0%; Pred. No. 9e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHOVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGOHCEIDKSKTCYEGNGHFRG 60  
Db 21 SNELHOVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGOHCEIDKSKTCYEGNGHFRG 80

QY 61 KASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNKCRNPDRRPPWCYVQVGLK 120  
Db 81 KASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNKCRNPDRRPPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155

```
RESULT 18
US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match          99.7%; Score 791; DB 12; Length 431;
Best Local Similarity 99.3%; Pred. No. 1.4e-71;
Matches 134; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
DB 141 PLVQECMVHDCADGK 155

RESULT 19
US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
; FILE REFERENCE: UTSN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-2

Query Match          98.7%; Score 783; DB 12; Length 411;
Best Local Similarity 99.3%; Pred. No. 8.7e-71;
Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60

RESULT 20
US-09-880-503-1
; Sequence 1, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-1

Query Match          64.3%; Score 510; DB 10; Length 88;
Best Local Similarity 100.0%; Pred. No. 5.2e-44;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 KTCYEGNGHFYRGKASTDTMTGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 107
DB 1 KTCYEGNGHFYRGKASTDTMTGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 60
QY 108 RRRPPWCYVQVGLKPLVQECMVHDCADGK 135
DB 61 RRRPPWCYVQVGLKPLVQECMVHDCADGK 88

RESULT 21
US-09-880-503-9
; Sequence 9, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-9

Query Match          64.3%; Score 510; DB 10; Length 96;
Best Local Similarity 100.0%; Pred. No. 5.7e-44;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
QY 48 KTCYEGNGHFYRKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPON 107
Db 1 KTCYEGNGHFYRKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPON 60
QY 108 RRRPCWYVQVGLKPLVQECMVHDCADGK 135
Db 61 RRRPCWYVQVGLKPLVQECMVHDCADGK 88

RESULT 22
US-09-987-457-18
; Sequence 18, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.218001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; TYPE: PRT
; ORGANISM: Homo sapiens (tpa)
US-09-987-457-18

Query Match 41.4%; Score 328.5; DB 11; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.7e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQQAALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 100
QY 59 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVG 118
Db 101 RGTWSTAESGAECTNNWSSALQKPYSGRRPDAILGLGNHNYCRNPDRDRSKWCYVFKA 160
QY 119 LKPLVQECMVHDCADG 134
Db 161 GKYSEFCSTPACSEG 176

RESULT 23
US-09-987-455-19
; Sequence 19, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.219001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1

QY 48 KTCYEGNGHFYRKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPON 107
Db 1 KTCYEGNGHFYRKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPON 60
QY 108 RRRPCWYVQVGLKPLVQECMVHDCADGK 135
Db 61 RRRPCWYVQVGLKPLVQECMVHDCADGK 88

RESULT 22
US-09-987-457-18
; Sequence 18, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.218001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; TYPE: PRT
; ORGANISM: Homo sapiens (tpa)
US-09-987-457-18

Query Match 41.4%; Score 328.5; DB 11; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.7e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQQAALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 100
QY 59 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVG 118
Db 101 RGTWSTAESGAECTNNWSSALQKPYSGRRPDAILGLGNHNYCRNPDRDRSKWCYVFKA 160
QY 119 LKPLVQECMVHDCADG 134
Db 161 GKYSEFCSTPACSEG 176

RESULT 23
US-09-987-455-19
; Sequence 19, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.219001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1

QY 48 KTCYEGNGHFYRKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPON 107
Db 1 KTCYEGNGHFYRKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPON 60
QY 108 RRRPCWYVQVGLKPLVQECMVHDCADGK 135
Db 61 RRRPCWYVQVGLKPLVQECMVHDCADGK 88

RESULT 22
US-09-987-457-18
; Sequence 18, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.218001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; TYPE: PRT
; ORGANISM: Homo sapiens (tpa)
US-09-987-457-18

Query Match 41.4%; Score 328.5; DB 11; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.7e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQQAALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 100
QY 59 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVG 118
Db 101 RGTWSTAESGAECTNNWSSALQKPYSGRRPDAILGLGNHNYCRNPDRDRSKWCYVFKA 160
QY 119 LKPLVQECMVHDCADG 134
Db 161 GKYSEFCSTPACSEG 176

RESULT 23
US-09-987-455-19
; Sequence 19, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.219001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 19
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-455-19

Query Match 41.4%; Score 328.5; DB 11; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.7e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQQAALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 100
QY 59 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVG 118
Db 101 RGTWSTAESGAECTNNWSSALQKPYSGRRPDAILGLGNHNYCRNPDRDRSKWCYVFKA 160
QY 119 LKPLVQECMVHDCADG 134
Db 161 GKYSEFCSTPACSEG 176

RESULT 24
US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-271-7

Query Match 41.4%; Score 328.5; DB 9; Length 562;
Best Local Similarity 46.3%; Pred. No. 7.2e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSEPRCFNGGTCQQAALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 135
QY 59 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVG 118
Db 136 RGTWSTAESGAECTNNWSSALQKPYSGRRPDAILGLGNHNYCRNPDRDRSKWCYVFKA 195
QY 119 LKPLVQECMVHDCADG 134
Db 196 GKYSEFCSTPACSEG 211

RESULT 25
US-09-974-298-145
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
```



```

; REFERENCE/DOCKET NUMBER: ANG-001/DIV (80144-0007)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 955-8787
; TELEFAX: (202) 955-3751
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Disulfide-bond
; LOCATION: group(11..19, 13..31, 33..42)
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-349-543-5

```

```

Query Match          36.9%; Score 293; DB 12; Length 49;
Best Local Similarity 100.0%; Pred. No. 2e-22;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFKFGGQHCEIDSKT 49
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFKFGGQHCEIDSKT 49

```

```

RESULT 29
US-09-880-503-7
; Sequence 7, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-7

```

```

Query Match          36.3%; Score 288; DB 10; Length 323;
Best Local Similarity 100.0%; Pred. No. 4.8e-21;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFKFGGQHCEIDSK 48
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFKFGGQHCEIDSK 48

```

```

RESULT 30
US-09-880-503-2
; Sequence 2, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2

```

```

; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-2

```

```

Query Match          35.7%; Score 283; DB 10; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.9e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFKFGGQHCEIDKS 47
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFKFGGQHCEIDKS 47

```

```

Search completed: December 3, 2003, 15:05:44
Job time : 14.8597 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:35:43 ; Search time 4.55172 Seconds  
(without alignments)  
818.010 Million cell updates/sec

Title: US-09-880-503-1

Perfect score: 510

Sequence: 1 KTCYEGNGHFRGKASTDTM.....QVGLKPLVQECMVHDCADGK 88

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	510	100.0	138	2	US-08-797-689-12
2	510	100.0	200	4	US-08-101-272G-73
3	510	100.0	208	4	US-08-101-272G-98
4	510	100.0	365	1	US-08-093-741-83
5	510	100.0	365	1	US-08-720-012-83
6	510	100.0	393	2	US-08-560-098A-44
7	510	100.0	393	3	US-08-967-024C-24
8	510	100.0	393	3	US-08-967-024C-25
9	510	100.0	411	1	US-08-087-163-1
10	510	100.0	411	1	US-08-286-748B-18
11	510	100.0	411	1	US-08-153-799-18
12	510	100.0	430	1	US-07-942-157A-3
13	510	100.0	430	6	5219569-2
14	510	100.0	431	4	US-08-101-272G-1
15	510	100.0	431	6	518829-1
16	510	100.0	432	2	US-08-560-098A-47
17	505	99.0	194	4	US-08-101-272G-80
18	505	99.0	201	4	US-08-101-272G-96
19	500	98.0	411	3	US-08-181-816-1
20	499	97.8	411	2	US-08-560-098A-48
21	489	95.9	89	4	US-08-101-272G-62
22	486	95.3	157	3	US-08-142-590B-25
23	241	47.3	477	2	US-08-560-098A-51
24	226	44.3	527	1	US-07-609-510B-16
25	226	44.3	527	2	US-08-811-949-39
26	226	44.3	527	5	PCT-US91-01025A-2
27	226	44.3	527	6	5185259-8

28	226	44.3	527	6	5520913-1
29	226	44.3	546	6	5200340-6
30	226	44.3	562	2	US-08-811-949-43
31	226	44.3	562	2	US-08-560-098A-50
32	226	44.3	562	2	US-08-883-795A-38
33	226	44.3	562	6	5185259-3
34	226	44.3	562	6	5200340-2
35	226	44.3	562	6	5344773-2
36	221	43.3	437	2	US-08-811-949-49
37	221	43.3	437	2	US-08-811-949-51
38	221	43.3	437	2	US-08-811-949-55
39	221	43.3	437	2	US-08-811-949-57
40	218	42.7	356	1	US-08-427-640-8
41	218	42.7	356	6	5244676-5
42	217	42.5	378	4	US-09-553-498-10
43	217	42.5	378	4	US-09-618-869-10
44	216	42.4	472	2	US-08-811-949-63
45	215	42.2	326	4	US-09-411-977-3

#### ALIGNMENTS

##### RESULT 1

US-08-797-689-12  
; Sequence 12, Application US/08797689  
; Patent No. 5875969  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; APPLICANT: Fournier, Alain  
; APPLICANT: Guitton, Jean-Dominique  
; APPLICANT: Jung, Gerard  
; APPLICANT: Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D., Julie K.  
; REGISTRATION NUMBER: P-38,619  
; REFERENCE/DOCKET NUMBER: ST92006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 138 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
US-08-797-689-12

Query Match 100.0%; Score 510; DB 2; Length 138;  
Best Local Similarity 100.0%; Pred. No. 8.5e-54;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 51 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 110  
Qy 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88  
Db 111 RRRPCYVQVGLKPLVQECMVHDCADGK 138

RESULT 2  
US-09-101-272G-73  
; Sequence 73, Application US/09101272G  
; Patent No. 6509445  
; GENERAL INFORMATION:  
; APPLICANT: Nissin Food Products Co., Ltd.  
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
; FILE REFERENCE: Q50979  
; CURRENT APPLICATION NUMBER: US/09/101,272G  
; CURRENT FILING DATE: 1998-07-08  
; PRIOR APPLICATION NUMBER: JP 1059/1996  
; PRIOR FILING DATE: 1996-01-08  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 73  
; LENGTH: 200  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: ATF domain of uPA  
US-09-101-272G-73

Query Match 100.0%; Score 510; DB 4; Length 200;  
Best Local Similarity 100.0%; Pred. No. 1.3e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127  
Qy 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88  
Db 128 RRRPCYVQVGLKPLVQECMVHDCADGK 155

RESULT 3  
US-09-101-272G-98  
; Sequence 98, Application US/09101272G  
; Patent No. 6509445  
; GENERAL INFORMATION:  
; APPLICANT: Nissin Food Products Co., Ltd.  
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
; FILE REFERENCE: Q50979  
; CURRENT APPLICATION NUMBER: US/09/101,272G  
; CURRENT FILING DATE: 1998-07-08  
; PRIOR APPLICATION NUMBER: JP 1059/1996  
; PRIOR FILING DATE: 1996-01-08  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 98  
; LENGTH: 208  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: ATFHI-ML chimeric protein  
US-09-101-272G-98

Query Match 100.0%; Score 510; DB 4; Length 208;  
Best Local Similarity 100.0%; Pred. No. 1.4e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 49 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 108  
Qy 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88  
Db 109 RRRPCYVQVGLKPLVQECMVHDCADGK 136

RESULT 4  
US-08-093-741-83  
; Sequence 83, Application US/08093741  
; Patent No. 5681721  
; GENERAL INFORMATION:  
; APPLICANT: STEFFENS, GERD J.  
; APPLICANT: WENDT, STEPHAN  
; APPLICANT: SCHNEIDER, JOHANNES  
; APPLICANT: HEINZEL-WIELAND, REGINA  
; APPLICANT: SAUNDERS, DEREK J.  
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH  
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN  
; TITLE OF INVENTION: INHIBITING EFFECT  
; NUMBER OF SEQUENCES: 83  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Evenson, Mckeown, Edwards & Lenahan  
; STREET: 1200 G Street, N. W. Suite 700  
; CITY: Washington, D.C.  
; COUNTRY: U.S.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/093.741  
; FILING DATE: 20-JUL-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P43 23 754.1  
; FILING DATE: 15-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: EVANS, Joseph D.  
; REGISTRATION NUMBER: 26,269  
; REFERENCE/DOCKET NUMBER: 148/41345  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)628-8800  
; TELEFAX: (202)628-8844  
; INFORMATION FOR SEQ ID NO: 83:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 365 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-093-741-83

Query Match 100.0%; Score 510; DB 1; Length 365;  
Best Local Similarity 100.0%; Pred. No. 2.7e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61  
Qy 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88  
Db 62 RRRPCYVQVGLKPLVQECMVHDCADGK 89

```
RESULT 5
US-08-720-012-83
; Sequence 83, Application US/08720012
; Patent No. 5747291
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, Gerd J.
; APPLICANT: WENNDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: SAUNDERS, DEREK J.
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
; TITLE OF INVENTION: INHIBITING EFFECT
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N. W. Suite 700
; CITY: Washington, D.C.
; COUNTRY: U.S.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720.012
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/093,741
; FILING DATE: 20-JUL-1993
; APPLICATION NUMBER: DE P43 23 754.1
; FILING DATE: 15-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-720-012-83
Query Match 100.0%; Score 510; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 2.7e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 61

Qy 61 RRRPCWYVQVGLKPLVQECMVHDCADGK 88
Db 62 RRRPCWYVQVGLKPLVQECMVHDCADGK 89

RESULT 6
US-08-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560.098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-44
Query Match 100.0%; Score 510; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 2.9e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60
Db 3 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 62

Qy 61 RRRPCWYVQVGLKPLVQECMVHDCADGK 88
Db 63 RRRPCWYVQVGLKPLVQECMVHDCADGK 90

RESULT 7
US-08-967-024C-24
; Sequence 24, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967.024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

06:30:55 2003

us-09-880-503-1.ra1

AN NUMBER: P 44 42 665.8  
RE: 30-NOV-1994  
ENT INFORMATION:  
ANS, Joseph D.  
ION NUMBER: 26,269  
E/DOCKET NUMBER: 148/42444  
ICATION INFORMATION:  
4E: (202) 628-8800  
: (202) 628-8844  
: 628-8844  
CHARACTERISTICS:  
: 393 amino acids  
: amino acid  
EDNESS:  
: linear  
s TYPE: protein  
4C-24

-h 100.0%; Score 510; DB 3; Length 393;  
1 Similarity 100.0%; Pred. No. 2.9e-53;  
88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHYCNRPDN 60  
3 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHYCNRPDN 62

61 RRRPCYVQVGLKPLVQECWHDGADGK 88  
63 RRRPCYVQVGLKPLVQECWHDGADGK 90

57-024C-25  
us 25 Application US/08967024C  
No 61301  
GENERAL INFORMATION:  
APPLICANT: WENDT, Stephan  
APPLICANT: STEFFENS, Gerd Josef  
APPLICANT: JANOSHA, Elke  
APPLICANT: HEINZEL-WIELAND, Regina  
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
STREET: 1200 G Street, N.W., Suite 700  
CITY: Washington  
STATE: DC  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/967,024C  
FILING DATE:  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: P 44 42 665.8  
FILING DATE: 30-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/42444  
TELEPHONE: (202) 628-8800  
TELEFAX: (202) 628-8844  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 393 amino acids  
TYPE: amino acid  
STRANDEDNESS:

MOLECULE TYPE: protein  
US-08-967-024C-25

Query Match 100.0%; Score 510; DB 3; Length 393;  
Best Local Similarity 100.0%; Pred. No. 2.9e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHYCNRPDN 60  
DB 3 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHYCNRPDN 62

QY 61 RRRPCYVQVGLKPLVQECWHDGADGK 88  
DB 63 RRRPCYVQVGLKPLVQECWHDGADGK 90

RESULT 9  
US-08-087-163-1  
Sequence 1, Application US/08087163  
Patent No. 5472692  
GENERAL INFORMATION:  
APPLICANT: Liu, Jian-Ning  
APPLICANT: Gurewicz, Victor  
TITLE OF INVENTION: PRO-UKINASE MUTANTS  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or SSX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: Patent In Release #1.0, Version 5.11  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/087,163  
FILING DATE: 07/02/93  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04353/003001  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 411  
TYPE: amino acid  
STRANDEDNESS: N/A  
TOPOLOGY: N/A

US-08-087-163-1

Query Match 100.0%; Score 510; DB 1; Length 411;  
Best Local Similarity 100.0%; Pred. No. 3.1e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHYCNRPDN 60  
DB 48 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLTQTYHAHRSALQGLGKHYCNRPDN 107

QY 61 RRRPCYVQVGLKPLVQECWHDGADGK 88  
DB 108 RRRPCYVQVGLKPLVQECWHDGADGK 135

## RESULT 10

US-08-286-748B-18  
; Sequence 18, Application US/08286748B  
; Patent No. 5759542  
; GENERAL INFORMATION:  
; APPLICANT: Victor Gurewlich  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY  
; TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF  
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 50Z or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,748B  
; FILING DATE: August 5, 1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: J. Peter Fasse  
; REGISTRATION NUMBER: 32,983  
; REFERENCE/DOCKET NUMBER: 04547/013001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 411  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

## US-08-286-748B-18

Query Match 100.0%; Score 510; DB 1; Length 411;  
Best Local Similarity 100.0%; Pred. No. 3.1e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 48 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107  
QY 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88  
Db 108 RRRPCYVQVGLKPLVQECMVHDCADGK 135  
RESULT 11  
US-08-153-799-18  
; Sequence 18, Application US/08153799  
; Patent No. 5766883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: R Hain Swope, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA

## ZIP: 07974

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB90/00650  
; FILING DATE: 26-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/775952  
; FILING DATE: 29-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Swope, R Hain  
; REGISTRATION NUMBER: 24864  
; REFERENCE/DOCKET NUMBER: 92H832  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 665 2400  
; TELEFAX: (908) 771 6159  
; TELEX: 219484  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 411 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-153-799-18

Query Match 100.0%; Score 510; DB 1; Length 411;  
Best Local Similarity 100.0%; Pred. No. 3.1e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 48 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107  
QY 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88  
Db 108 RRRPCYVQVGLKPLVQECMVHDCADGK 135

## RESULT 12

US-07-942-157A-3  
; Sequence 3, Application US/07942157A  
; Patent No. 5648253  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Cha-Mer  
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/942,157A



Thu Dec 4 06:30:55 2003

```
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TS1108Cont.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)815-6508
; TELEFAX: (404)815-6555
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..19
; OTHER INFORMATION: /label= peptide
; OTHER INFORMATION: /note= "WAP signal"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 198..203
; OTHER INFORMATION: /label= modified
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"
; US-07-942-157A-3

Query Match 100.0%; Score 510; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 3.3e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 67 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 126
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 127 RRRPWCYVQVGLKPLVQECMVHDCADGK 154

RESULT 13
5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO: 2:
; LENGTH: 430
5219569-2

Query Match 100.0%; Score 510; DB 6; Length 430;
Best Local Similarity 100.0%; Pred. No. 3.3e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155
```

```
RESULT 14
US-09-101-272G-1
; Sequence 1, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1:
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (21)..()
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (20)..()
; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
; US-09-101-272G-1

Query Match 100.0%; Score 510; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.3e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 15
5188829-1
; Patent No. 5188829
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIKAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1988
; SEQ ID NO: 1:
; LENGTH: 431
5188829-1

Query Match 100.0%; Score 510; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.3e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 16
US-08-560-098A-47
; Sequence 47, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNENDT, Stephan
```

APPLICANT: HEINZEL-WIELAND, Regina  
APPLICANT: STEFFENS, Gerd Josef  
TITLE OF INVENTION: Proteins having Fibrinolytic and  
TITLE OF INVENTION: Coagulation-inhibiting Properties  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
STREET: 1200 G Street, N.W., Suite 700  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/560,098A  
FILING DATE: 17-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: P 44 40 892.7  
FILING DATE: 17-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/42448  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-8800  
TELEFAX: (202) 628-8844  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 432 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-560-098A-47

Query Match 100.0%; Score 510; DB 2; Length 432;  
Best Local Similarity 100.0%; Pred. No. 3.3e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 60  
Db 69 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 128  
QY 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 129 RRPWCYVQVGLKPLVQECMVHDCADGK 156

RESULT 17  
US-09-101-272G-80  
Sequence 80, Application US/09101272G  
Patent No. 6509445  
GENERAL INFORMATION:  
APPLICANT: Nissin Food Products Co., Ltd.  
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
FILE REFERENCE: Q50979  
CURRENT APPLICATION NUMBER: US/09/101,272G  
CURRENT FILING DATE: 1998-07-08  
PRIOR APPLICATION NUMBER: JP 1059/1996  
PRIOR FILING DATE: 1996-01-08  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 80  
LENGTH: 194  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: ATFHI chimeric protein  
US-09-101-272G-80

Query Match 99.0%; Score 505; DB 4; Length 194;  
Best Local Similarity 100.0%; Pred. No. 5.1e-53;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 60  
Db 49 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 108  
QY 61 RRPWCYVQVGLKPLVQECMVHDCADG 87  
Db 109 RRPWCYVQVGLKPLVQECMVHDCADG 135

RESULT 18  
US-09-101-272G-96  
Sequence 96, Application US/09101272G  
Patent No. 6509445  
GENERAL INFORMATION:  
APPLICANT: Nissin Food Products Co., Ltd.  
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
FILE REFERENCE: Q50979  
CURRENT APPLICATION NUMBER: US/09/101,272G  
CURRENT FILING DATE: 1998-07-08  
PRIOR APPLICATION NUMBER: JP 1059/1996  
PRIOR FILING DATE: 1996-01-08  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 96  
LENGTH: 201  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: ATFHI-CL chimeric protein  
US-09-101-272G-96

Query Match 99.0%; Score 505; DB 4; Length 201;  
Best Local Similarity 100.0%; Pred. No. 5.3e-53;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 60  
Db 49 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 108  
QY 61 RRPWCYVQVGLKPLVQECMVHDCADG 87  
Db 109 RRPWCYVQVGLKPLVQECMVHDCADG 135

RESULT 19  
US-09-181-816-1  
Sequence 1, Application US/09181816  
Patent No. 6277818  
GENERAL INFORMATION:  
APPLICANT: MAZAR, Andrew P.  
APPLICANT: JONES, Terence R.  
TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE  
TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR  
FILE REFERENCE: 329042000300 SIDN 1-7  
CURRENT APPLICATION NUMBER: US/09/181,816  
CURRENT FILING DATE: 1998-10-29  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 411  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-181-816-1

Query Match 98.0%; Score 500; DB 3; Length 411;  
Best Local Similarity 98.9%; Pred. No. 4.9e-52;  
Matches 87; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query Match 95.3%; Score 486; DB 3; Length 157;

```
; Best Local Similarity 95.5%; Pred. No. 7.5e-51;
Matches 84; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
D 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 107
QY 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88
D 108 RRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 23
US-08-560-098A-51
; Sequence 51, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-51
Query Match 47.3%; Score 241; DB 2; Length 477;
Best Local Similarity 50.0%; Pred. No. 7.4e-21;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 61
D 127 TCYKQGVYRGVTWSTESGAQCNWNSNLLTRTYNGRSDAITLGLGNHNYCRNPDN 186
QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
D 187 SKPWCYVIRKASKFILEFCVPVCS 210

RESULT 24
US-07-609-510B-16
; Sequence 16, Application US/07609510B
```

```
; Patent No. 5326700
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue F
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,510B
; FILING DATE: 19901106
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-609-510B-16
Query Match 44.3%; Score 226; DB 1; Length 527;
Best Local Similarity 47.7%; Pred. No. 5.3e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 61
D 91 TCYEDQGISYRGVTWSTESGAECTNWSALAKPKYSGRRPDAIRLGLGNHNYCRNPD 150
QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
D 151 SKPWCYVFKAGYSFECSTPACSEG 176

RESULT 25
US-08-811-949-39
; Sequence 39, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:43:04 ; Search time 9.03448 Seconds  
(without alignments)  
1811.566 Million cell updates/sec

Title: US-09-880-503-1  
Perfect score: 510  
Sequence: 1 KTCYEGNGHFGYRGKASTDTM.....QVGLKPLVQECMVHDCADGK 88

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues  
Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	510	100.0	88	US-09-880-503-1	Sequence 1, Appli
2	510	100.0	96	US-09-880-503-9	Sequence 9, Appli
3	510	100.0	135	US-09-880-503-4	Sequence 4, Appli
4	510	100.0	138	US-09-884-186-12	Sequence 12, Appli
5	510	100.0	138	US-10-237-667-12	Sequence 12, Appli
6	510	100.0	138	US-10-237-708-12	Sequence 12, Appli
7	510	100.0	138	US-10-237-866-12	Sequence 12, Appli
8	510	100.0	138	US-10-237-871-12	Sequence 12, Appli
9	510	100.0	138	US-10-237-624-12	Sequence 12, Appli
10	510	100.0	143	US-09-880-503-8	Sequence 8, Appli
11	510	100.0	337	US-10-106-698-6266	Sequence 6266, Ap
12	510	100.0	403	US-09-880-503-6	Sequence 6, Appli
13	510	100.0	411	US-09-880-503-3	Sequence 3, Appli
14	510	100.0	431	US-09-264-468B-1	Sequence 1, Appli
15	510	100.0	431	US-10-301-822-161	Sequence 161, App

16	510	100.0	431	12	US-10-131-985-21	Sequence 21, Appli
17	510	100.0	431	14	US-10-076-421-2	Sequence 2, Appli
18	510	100.0	431	15	US-10-171-311-184	Sequence 184, App
19	510	100.0	431	15	US-10-193-656-4	Sequence 4, Appli
20	508	98.6	431	12	US-10-247-671-149	Sequence 149, App
21	500	98.0	411	12	US-10-407-821-2	Sequence 2, Appli
22	226	44.3	527	11	US-09-987-451-18	Sequence 18, Appli
23	226	44.3	527	11	US-09-987-455-19	Sequence 19, Appli
24	226	44.3	562	9	US-09-969-271-7	Sequence 7, Appli
25	226	44.3	562	10	US-09-974-298-145	Sequence 145, App
26	226	44.3	562	12	US-10-443-701-4	Sequence 4, Appli
27	226	44.3	562	15	US-10-193-656-8	Sequence 8, Appli
28	215	42.2	326	14	US-10-057-951-3	Sequence 3, Appli
29	215	42.2	354	11	US-09-987-457-10	Sequence 10, Appli
30	215	42.2	354	11	US-09-987-455-11	Sequence 11, Appli
31	215	42.2	372	9	US-09-084-491A-3	Sequence 3, Appli
32	215	42.2	372	14	US-10-102-704-3	Sequence 3, Appli
33	215	42.2	377	11	US-09-987-455-8	Sequence 8, Appli
34	193.5	37.9	560	10	US-09-912-559-3	Sequence 3, Appli
35	193.5	37.9	560	10	US-09-912-559-4	Sequence 4, Appli
36	193.5	37.9	560	15	US-10-172-712-32	Sequence 32, Appli
37	193	37.8	655	15	US-10-172-712-28	Sequence 28, Appli
38	176	34.5	343	11	US-09-987-457-14	Sequence 14, Appli
39	176	34.5	343	11	US-09-987-455-15	Sequence 15, Appli
40	175	34.3	339	11	US-09-987-457-12	Sequence 12, Appli
41	175	34.3	339	11	US-09-987-455-13	Sequence 13, Appli
42	173	33.9	83	12	US-09-981-151A-86	Sequence 86, Appli
43	172	33.7	331	11	US-09-987-457-11	Sequence 11, Appli
44	172	33.7	331	11	US-09-987-455-12	Sequence 12, Appli
45	171	33.5	79	12	US-09-981-151A-85	Sequence 85, Appli

ALIGNMENTS

RESULT 1

US-09-880-503-1  
; Sequence 1, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; TITLE OF INVENTION: TISSUE CONTRACTILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 88  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-1

Query Match 100.0%; Score 510; DB 10; Length 88;  
Best Local Similarity 100.0%; Pred. No. 3.5e-52;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KTCYEGNGHFGYRGKASTDTMGRPCLPWSNGATVLCQYTHAHRSDALQGLGKHNYCRNPDN 60
Db	1	KTCYEGNGHFGYRGKASTDTMGRPCLPWSNGATVLCQYTHAHRSDALQGLGKHNYCRNPDN 60
QY	61	RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db	61	RRRPWCYVQVGLKPLVQECMVHDCADGK 88

RESULT 2

US-09-880-503-9  
; Sequence 9, Application US/09880503

```
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-880-503-9

Query Match      100.0%; Score 510; DB 10; Length 96;
Best Local Similarity 100.0%; Pred. No. 3.9e-52;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNP DN 60
Db      1 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNP DN 60

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88

RESULT 3
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-880-503-4

Query Match      100.0%; Score 510; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 5.6e-52;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNP DN 60
Db      48 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNP DN 107

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 4
US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

Query Match      100.0%; Score 510; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.7e-52;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNP DN 60
Db      51 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNP DN 110

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 5
US-10-237-667-12
; Sequence 12, Application US/10237667
; Publication No. US20030022308A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
```

PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (Patentin)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,667  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-237-667-12

Query Match 100.0%; Score 510; DB 15; Length 138;  
Best Local Similarity 100.0%; Pred. No. 5.7e-52;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
Db 51 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 110

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 6  
US-10-237-708-12  
Sequence 12, Application US/10237708  
Publication No. US20030036170A1  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guitton, Jean-Dominique  
Jung, Gerard  
Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426

STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (Patentin)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,708  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-237-708-12

Query Match 100.0%; Score 510; DB 15; Length 138;  
Best Local Similarity 100.0%; Pred. No. 5.7e-52;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
Db 51 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 110

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 7  
US-10-237-866-12  
Sequence 12, Application US/10237866  
Publication No. US20030036171A1  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guitton, Jean-Dominique  
Jung, Gerard  
Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,866  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-237-866-12  
Query Match 100.0%; Score 510; DB 15; Length 138;  
Best Local Similarity 100.0%; Pred. No. 5.7e-52;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KTCYEGNGHFRYGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60  
Db 51 KTCYEGNGHFRYGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 110  
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138  
RESULT 8  
US-10-237-871-12  
Sequence 12, Application US/10237871  
Publication No. US20030036172A1  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guitton, Jean-Dominique  
Jung, Gerard  
Yeh, Patrice  
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,871  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-237-871-12  
Query Match 100.0%; Score 510; DB 15; Length 138;  
Best Local Similarity 100.0%; Pred. No. 5.7e-52;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KTCYEGNGHFRYGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60  
Db 51 KTCYEGNGHFRYGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 110  
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138  
RESULT 9  
US-10-237-624-12  
Sequence 12, Application US/10237624  
Publication No. US20030082747A1  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guitton, Jean-Dominique  
Jung, Gerard  
Yeh, Patrice  
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,624  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:



```
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3829
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-624-12

Query Match 100.0%; Score 510; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.7e-52;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60
Db 51 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 110

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 10
US-09-880-503-8
; Sequence 8, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-8

Query Match 100.0%; Score 510; DB 10; Length 143;
Best Local Similarity 100.0%; Pred. No. 5.9e-52;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 11
```

```
US-10-106-698-6266
; Sequence 6266, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypepti
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6266
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-6266

Query Match 100.0%; Score 510; DB 15; Length 337;
Best Local Similarity 100.0%; Pred. No. 1.5e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60
Db 74 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 133

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 134 RRRPWCYVQVGLKPLVQECMVHDCADGK 161

RESULT 12
US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-6

Query Match 100.0%; Score 510; DB 10; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.8e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFGYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 13
US-09-880-503-3
```

```
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-3

Query Match      100.0%; Score 510; DB 10; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db      48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 14
US-09-264-468B-1
; Sequence 1, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310-US.P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
US-09-264-468B-1

Query Match      100.0%; Score 510; DB 10; Length 431;
Best Local Similarity 100.0%; Pred. No. 2e-51;

Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db      48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

US-09-880-503-1

Query Match      100.0%; Score 510; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 2e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db      68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 15
US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Subhangai
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: BURGART, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-0292P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match      100.0%; Score 510; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 2e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db      68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 16
US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: Mcintosh, Fraser S
; APPLICANT: Occleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
```

```
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-21

Query Match
Best Local Similarity 100.0%; Score 510; DB 12; Length 431;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 17
US-10-076-421-2
; Sequence 2, Application US/10076421
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, NAKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAYAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-2

Query Match
Best Local Similarity 100.0%; Score 510; DB 14; Length 431;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 18
US-10-171-311-184
; Sequence 184, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
```

```
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-184

Query Match
Best Local Similarity 100.0%; Score 510; DB 15; Length 431;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 19
US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)..(431)
US-10-193-656-4

Query Match
Best Local Similarity 100.0%; Score 510; DB 15; Length 431;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155
```

Db 128 RRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 20

US-10-247-671-149  
; Sequence 149, Application US/10247671  
; Publication No. US20030194721A1  
; GENERAL INFORMATION:  
; APPLICANT: Mikita, Thomas  
; APPLICANT: Shiffman, Dov  
; APPLICANT: Porter, Gordon, J.  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS  
; FILE REFERENCE: PA-0050 US  
; CURRENT APPLICATION NUMBER: US/10/247,671  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/323,784  
; PRIOR FILING DATE: 2001-09-19  
; NUMBER OF SEQ ID NOS: 186  
; SOFTWARE: PERL Program  
; SEQ ID NO 149  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1  
; US-10-247-671-149

Query Match 99.6%; Score 508; DB 12; Length 431;  
Best Local Similarity 98.9%; Pred. No. 3.4e-51;  
Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 68 KTCYEGNGHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127

Qy 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 128 RRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 21

US-10-407-821-2  
; Sequence 2, Application US/10407821  
; Publication No. US20030219386A1  
; GENERAL INFORMATION:  
; APPLICANT: IDELL, STEVEN  
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED  
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL  
; FILE REFERENCE: UTSN:022US  
; CURRENT APPLICATION NUMBER: US/10/407,821  
; PRIOR FILING DATE: 2003-04-04  
; PRIOR APPLICATION NUMBER: 60/414,202  
; PRIOR FILING DATE: 2002-09-27  
; PRIOR APPLICATION NUMBER: 60/370,456  
; PRIOR FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-407-821-2

Query Match 98.0%; Score 500; DB 12; Length 411;  
Best Local Similarity 98.9%; Pred. No. 2.8e-50;  
Matches 87; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 48 KTCYEGNGHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107

Qy 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 108 RRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 22

US-09-987-457-18  
; Sequence 18, Application US/09987457  
; Publication No. US20030013150A1  
; GENERAL INFORMATION:  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Manosroi, Jiradej  
; APPLICANT: Tavapiwatana, Chatchai  
; APPLICANT: Goetz, Friedrich  
; APPLICANT: Werner, Rolf-Guenther  
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes  
; FILE REFERENCE: 0652.218001  
; CURRENT APPLICATION NUMBER: US/09/987,457  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,573  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 00 27 782.2  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Homo sapiens (tPA)  
; US-09-987-457-18

Query Match 44.3%; Score 226; DB 11; Length 527;  
Best Local Similarity 47.7%; Pred. No. 4.3e-19;  
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

Qy 2 TCYEGNGHFRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNR 61  
Db 91 TCYEDQGISYRGWTSTAESGAECTNNSSALAQPKSYGRRPDAIRLGLGNHNYCRNPDRO 150

Qy 62 RRPWCYVQVGLKPLVQECMVHDCADG 87  
Db 151 SKPCYVFKAGYSSEFCSTPACSEG 176

RESULT 23

US-09-987-455-19  
; Sequence 19, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tavapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.219001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-987-455-19

Query Match 44.3%; Score 226; DB 11; Length 527;

```
Best Local Similarity 47.7%; Pred. No. 4.3e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCPLPMNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDNR 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 TCYEDQGISYRGTWSTAESGAECTNWNSSALAAQKPYSGRRPDARLGLGNHNYCRNPD 185
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 91 TCYEDQGISYRGTWSTAESGAECTNWNSSALAAQKPYSGRRPDARLGLGNHNYCRNPD 150
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 151 SKPCYVYFKAGKYSSEFCSTPACSEG 176
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 24
US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-271-7

Query Match 44.3%; Score 226; DB 9; Length 562;
Best Local Similarity 47.7%; Pred. No. 4.6e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCPLPMNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDNR 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 TCYEDQGISYRGTWSTAESGAECTNWNSSALAAQKPYSGRRPDARLGLGNHNYCRNPD 185
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186 SKPCYVYFKAGKYSSEFCSTPACSEG 211
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 25
US-09-974-298-145
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 145
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145

Query Match 44.3%; Score 226; DB 10; Length 562;
Best Local Similarity 47.7%; Pred. No. 4.6e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCPLPMNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDNR 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 TCYEDQGISYRGTWSTAESGAECTNWNSSALAAQKPYSGRRPDARLGLGNHNYCRNPD 185
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186 SKPCYVYFKAGKYSSEFCSTPACSEG 211
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 26
US-10-443-701-4
; Sequence 4, Application US/10443701
; Publication No. US20030195016A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P1788R1
; CURRENT APPLICATION NUMBER: US/10/443,701
; CURRENT FILING DATE: 2003-05-21
; PRIOR APPLICATION NUMBER: US/09/703,695
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-443-701-4

Query Match 44.3%; Score 226; DB 12; Length 562;
Best Local Similarity 47.7%; Pred. No. 4.6e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCPLPMNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDNR 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 TCYEDQGISYRGTWSTAESGAECTNWNSSALAAQKPYSGRRPDARLGLGNHNYCRNPD 185
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186 SKPCYVYFKAGKYSSEFCSTPACSEG 211
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 27
US-10-193-656-8
; Sequence 8, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00750
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)..(562)
US-10-193-656-8
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:58 ; Search time 4.55172 Seconds  
(without alignments)  
1859.261 Million cell updates/sec

Title: US-09-880-503-1  
Perfect score: 510  
Sequence: 1 KTCYEGNGHFRGKASTDTM.....QVGLKPLVQECMVHDCADGK 88

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76:\*  
1: Pir1:\*  
2: Pir2:\*  
3: Pir3:\*  
4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	431	1 UKHU	u-plasminogen acti
2	475	93.1	433	1 UKBAY	u-plasminogen acti
3	420	82.4	442	1 UKPG	u-plasminogen acti
4	406	79.6	432	1 S18932	u-plasminogen acti
5	396	77.6	433	1 JN0560	u-plasminogen acti
6	387	75.9	433	1 UKMS	u-plasminogen acti
7	241	47.3	431	2 JS0599	t-plasminogen acti
8	241	47.3	477	1 A34369	t-plasminogen acti
9	241	47.3	477	2 JS0598	t-plasminogen acti
10	226	44.3	291	2 JS0598	t-plasminogen acti
11	226	44.3	562	1 UKHUT	t-plasminogen acti
12	225	44.1	559	1 A35029	t-plasminogen acti
13	220	43.1	394	2 JS0600	t-plasminogen acti
14	219	42.9	559	1 A29941	t-plasminogen acti
15	213	41.8	477	2 JS0597	t-plasminogen acti
16	209.5	41.1	558	2 JC5878	plasma hyaluronan-
17	198.5	38.9	434	1 A35005	plasma hyaluronan-
18	193.5	37.9	560	1 JC4795	plasma hyaluronan-
19	193	37.8	655	1 A46688	hepatocyte growth
20	180.5	35.4	603	2 S28841	coagulation factor
21	170.5	33.4	615	1 KFHU12	coagulation factor
22	161	31.6	1420	2 A32869	apolipoprotein(a)
23	156	30.6	1420	2 E61545	plasmin (EC 3.4.21
24	156	30.6	4548	1 S00657	apoptein(a) (EC
25	155.5	30.5	593	2 S45281	coagulation factor
26	153	30.0	89	2 A60140	plasmin (EC 3.4.21
27	150	29.4	123	2 G61545	plasmin (EC 3.4.21
28	149.5	29.3	169	2 A40522	plasmin (EC 3.4.21
29	149	29.2	812	1 PLBO	plasmin (EC 3.4.21

30	148.5	29.1	810	2	I46260	plasmin (EC 3.4.21
31	148	29.0	790	1	PLPG	plasmin (EC 3.4.21
32	147.5	28.9	460	2	B61545	plasmin (EC 3.4.21
33	147.5	28.9	943	2	B45082	neurotrophic recep
34	146	28.6	937	2	A45082	neurotrophic recep
35	144	28.2	810	1	PLHU	plasmin (EC 3.4.21
36	143	28.0	806	2	T18840	hypothetical prote
37	142.5	27.9	711	1	A47136	macrophage-stimula
38	141	27.6	810	2	B30848	plasmin (EC 3.4.21
39	139	27.3	685	1	A48289	neurotrophic recep
40	137.5	27.0	728	1	A60185	hepatocyte growth
41	135.5	26.6	728	1	A35644	hepatocyte growth
42	134	26.3	455	2	A61545	plasmin (EC 3.4.21
43	133.5	26.2	728	1	JH0579	hepatocyte growth
44	132.5	26.0	411	2	I51285	hepatocyte growth
45	132.5	26.0	2869	2	T18518	apolipoprotein(a)

## ALIGNMENTS

## RESULT 1

UKHU  
u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human  
N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen a  
N;Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen a  
in form  
C;Species: Homo sapiens (man)  
C;Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000  
A;Accession: A00931; I52209; J70102; A37561; I38102; S65783; A37562; A37563; A37564; A  
R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastiao, G.; Boast, S.; Blasi, F.  
Nucleic Acids Res. 13, 2759-2771, 1985  
A;Title: The human urokinase-plasminogen activator gene and its promoter.  
A;Reference number: A00931; MUID:85215647; PMID:2987867  
A;Accession: A00931  
A;Molecule type: DNA  
A;Residues: 1-431 <RTC>  
A;Cross-references: GB:X02419; NID:G37601; PIDN:CAR26268.1; PID:G1834524  
A;Note: the authors translated the codon ATG for residue 214 as Ile  
R;Nagamine, Y.; Pearson, D.; Grattan, M.  
Biochem. Biophys. Res. Commun. 132, 563-569, 1985  
A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcin  
A;Reference number: I52209; MUID:86050639; PMID:3933505  
A;Accession: I52209  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 145-161 <NAG1>  
A;Cross-references: GB:X03027; NID:G340174; PIDN:AAA61257.1; PID:G340175  
R;Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama  
Gene 36, 183-188, 1985  
A;Title: Molecular cloning of cDNA coding for human preprourokinase.  
A;Reference number: J70102; MUID:86056954; PMID:2415429  
A;Accession: J70102  
A;Molecule type: mRNA  
A;Residues: 1-213, 'I', 215-431 <NAG2>  
A;Cross-references: GB:X03226; NID:G340155; PIDN:AA09138.1; PID:G340158; GB:D00244; N  
R;Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.  
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984  
A;Title: Identification and primary sequence of an unspliced human urokinase poly(A) +  
A;Reference number: A37561; MUID:84272706; PMID:6589620  
A;Accession: A37561  
A;Molecule type: mRNA  
A;Residues: 66-431 <VER>  
A;Cross-references: GB:D00244; NID:G220138  
R;Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elise  
DNA 4, 139-146, 1985  
A;Title: Molecular cloning, sequencing, and expression in Escherichia coli of human pr  
A;Reference number: I38102; MUID:85203359; PMID:3888571  
A;Accession: I38102  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>  
A;Cross-references: EMBL:X02760; NID:G35297; PIDN:CAA26535.1; PID:G35298

R; Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, K. *Biochim. Biophys. Acta* 1293, 88-89, 1996  
 A; Title: Characterization of single chain urokinase-type plasminogen activator with a novel structure.  
 A; Reference number: S65783; MUID:96186279; PMID:8652631  
 A; Accession: S65783  
 A; Status: preliminary  
 A; Molecule type: mRNA  
 A; Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>  
 A; Cross-references: EMBL:D11143; NID:G1311467; PIDN:BA01919.1; PID:G1199928  
 R; Gunzler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L.  
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982  
 A; Title: The primary structure of high molecular mass urokinase from human urine.  
 A; Reference number: A37562; MUID:83055084; PMID:6754569  
 A; Accession: A37562  
 A; Molecule type: protein  
 A; Residues: 21-177 <GUN>  
 R; Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O.  
 Eur. J. Biochem. 125, 251-257, 1982  
 A; Title: Human low-molecular-weight urinary urokinase. Partial characterization and preliminary structure.  
 A; Reference number: A37563; MUID:83003608; PMID:6749491  
 A; Accession: A37563  
 A; Molecule type: protein  
 A; Residues: 156-176; 179-193, 'T', 195, 'T', 197-224 <SCH>  
 R; Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Flohe, L.  
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982  
 A; Title: The complete amino acid sequence of low molecular mass urokinase from human urine.  
 A; Reference number: A37564; MUID:83055099; PMID:6754572  
 A; Accession: A37564  
 A; Molecule type: protein  
 A; Residues: 158-410 <STE>  
 R; Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K.  
 Biochem. Biophys. Res. Commun. 171, 401-406, 1990  
 A; Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinant urokinase.  
 A; Reference number: A35689; MUID:90365737; PMID:2393398  
 A; Accession: A35689  
 A; Molecule type: protein  
 A; Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>  
 A; Note: identification of a fucose and attempt to determine its attachment site  
 R; Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzman, I.  
 Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990  
 A; Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell line.  
 A; Reference number: A36697; MUID:91097529; PMID:2125213  
 A; Accession: A36697  
 A; Molecule type: protein  
 A; Residues: 21-34 <RAB>  
 R; Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.  
 Submitted to the Brookhaven Protein Data Bank, July 1993  
 A; Reference number: A51255; PDB:1KDU  
 A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue R; Li, X.; Smith, R.A.G.; Dobson, C.M.  
 Biochemistry 31, 9562-9571, 1992  
 A; Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain of urokinase.  
 A; Reference number: A44375; MUID:93003110; PMID:1327118  
 A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR  
 R; Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazar, A.P.; Olejniczak, Z.  
 Submitted to the Brookhaven Protein Data Bank, January 1994  
 A; Reference number: A66822; PDB:1URK  
 A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue R; Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.; submitted to the Brookhaven Protein Data Bank, July 1995  
 A; Reference number: A66058; PDB:1LMW  
 A; Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426  
 C; Comment: This enzyme is found in urine in a high molecular mass form, consisting of A and B subunits.  
 C; Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a zymogen of plasmin.  
 C; Genetics:  
 A; Gene: GDB:PLAU  
 A; Cross-references: GDB:119497; OMIM:191840  
 A; Map position: 10q24-10q24  
 A; Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3  
 C; Function:  
 A; Description: proteolytically activates plasminogen  
 A; Pathway: fibrinolysis  
 C; Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C; Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protease; urokinase-type plasminogen activator  
 F; 1-20/Domain: signal sequence #status predicted <SIG>  
 F; 21-431/Product: urokinase-type plasminogen activator, single chain form #status predicted <SIG>  
 F; 21-177/Product: urokinase-type plasminogen activator chain A #status experimental <M>  
 F; 31-62/Domain: EGF homology <EGF>  
 F; 70-151/Domain: kringle homology <KRG>  
 F; 156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental  
 F; 179-431/Product: urokinase-type plasminogen activator chain B #status experimental <KRG>  
 F; 179-419/Domain: trypsin homology <TRY>  
 F; 31-39, 33-51, 53-62, 70-151, 91-133, 122-146, 168-299, 209-225, 217-288, 313-382, 345-361, 372-382/Binding site: carbohydrate (Thr) (covalent) #status predicted  
 F; 38/Binding site: carbohydrate (Thr) (covalent) #status predicted  
 F; 178-179/Cleavage site: Lys-Ile (plasma) #status experimental  
 F; 224, 275, 376/Active site: His, Asp, Ser #status experimental  
 F; 322/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 Query Match 100.0%; Score 510; DB 1; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-48;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNSTATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60  
 Db |||||  
 Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 Db |||||  
 Qy 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155  
 Db |||||  
 RESULT 2  
 UKBAY  
 u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon  
 C; Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)  
 C; Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 18-Jun-1999  
 C; Accession: S14687; S08651  
 R; Au, Y.P.T.; Wang, T.W.; Clowes, A.W.  
 Nucleic Acids Res. 18, 3411, 1990  
 A; Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator.  
 A; Reference number: S14687; MUID:90287734; PMID:2113276  
 A; Accession: S14687  
 A; Molecule type: mRNA  
 A; Residues: 1-433 <AU>  
 A; Cross-references: EMBL:X51935; NID:G38130; PIDN:CAA36200.1; PID:G38131  
 C; Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t  
 C; Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F; 1-20/Domain: signal sequence #status predicted <SIG>  
 F; 21-176/Product: plasminogen activator chain A #status predicted <ACH>  
 F; 30-61/Domain: EGF homology <EGF>  
 F; 69-150/Domain: kringle homology <KRG>  
 F; 178-433/Product: plasminogen activator chain B #status predicted <BCH>  
 F; 178-421/Domain: trypsin homology <TRY>  
 F; 167-238, 208-224, 216-287, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted  
 F; 223, 274, 378/Active site: His, Asp, Ser #status predicted  
 F; 324/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 93.1%; Score 475; DB 1; Length 433;  
 Best Local Similarity 94.3%; Pred. No. 2.9e-44;  
 Matches 83; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNSTATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60  
 Db |||||  
 Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 Db |||||  
 Qy 127 RRRPWCYVQVGLKPLVQECMVHDCADGK 154  
 Db |||||  
 RESULT 3  
 UKPG  
 u-plasminogen activator (EC 3.4.21.73) precursor - pig  
 N; Alternate names: uPA  
 C; Species: Sus scrofa domestica (domestic pig)  
 C; Date: 04-Dec-1986 #sequence\_revision 17-Mar-1987 #text\_change 07-Aug-1998



A:Cross-references: EMBL:X66907; NID:g396200; PIDN:CAA47356.1; PID:g938279

C:Genetics: uPA

A:Gene: uPA

C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; tC

C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>

F:31-62/Domain: EGF homology <EGF>

F:70-151/Domain: kringle homology <KRG>

F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>

F:179-430/Domain: trypsin homology <TRY>

F:168-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status predicted

F:225,276,377/Active site: His, Asp, Ser #status predicted

Query Match 79.6%; Score 406; DB 1; Length 432;

Best Local Similarity 79.5%; Pred. No. 1e-36;

Matches 70; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFGKASTDTMGRCPLPWNSATVLQQTYHAHRSDALQLGLGKHKNYCRNPDN 60

Db 68 KTCYHNGQSYKRGKANTDTKGRPCLANSPAVLQQTYHAHRSDALSGLGKHKNYCRNPDN 127

Qy 61 RRRPWCYVOGKLPVQECMVHDCADGK 88

Db 128 QRRPWCYVOIGLKQFVQECMVQDCSLSK 155

RESULT 5

JN0560

U:Plasminogen activator (EC 3.4.21.73) precursor - bovine

N:Alternate names: uPA

C:Species: Bos primigenius taurus (cattle)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1999 #text\_change 16-Jul-1999

C:Accession: JN0560

R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.

Gene 125, 177-183, 1993

A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and

A:Reference number: JN0560; MUID:93216119; PMID:8385052

A:Accession: JN0560

A:Molecule type: mRNA

A:Residues: 1-433 <KRA>

A:Cross-references: GB:L03546; NID:gl63800; PIDN:AAA51419.1; PID:gl63801

C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t

C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <MA1>

F:33-64/Domain: EGF homology <EGF>

F:72-153/Domain: kringle homology <KRG>

F:181-433/Product: plasminogen activator chain B #status predicted <MA2>

F:181-421/Domain: trypsin homology <TRY>

F:170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted

F:226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 77.6%; Score 396; DB 1; Length 433;

Best Local Similarity 76.1%; Pred. No. 1.3e-35;

Matches 67; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFGKASTDTMGRCPLPWNSATVLQQTYHAHRSDALQLGLGKHKNYCRNPDN 60

Db 70 KTCYQNGHSGYRGKANRDLGRCPLAWDSPTVLLKMYHAHRSDAQLGLGKHKNYCRNPDN 129

Qy 61 RRRPWCYVOGKLPVQECMVHDCADGK 88

Db 130 QRRPWCYVOIGLKQFVQECMVQDCSVGK 157

RESULT 6

UKMS

U:Plasminogen activator (EC 3.4.21.73) precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 18-Jun-1999

C:Accession: A29420; A24615

R;Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.  
 Biochemistry 26, 8270-8279, 1987  
 A;Title: The murine urokinase-type plasminogen activator gene.  
 A;Reference number: A29420; MUID:88163489; PMID:2831940  
 A;Accession: A29420  
 A;Molecule type: DNA  
 A;Residues: 1-433 <DEG>  
 A;Cross-references: GB:M17922; NID:g202296; PIDN:AAA40539.1; PID:g202297  
 R;Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher, Eur. J. Biochem. 148, 225-232, 1985  
 A;Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase  
 A;Reference number: A24615; MUID:85179474; PMID:2985383  
 A;Accession: A24615  
 A;Molecule type: mRNA  
 A;Residues: 1-433 <BEL>  
 A;Cross-references: GB:X02389; NID:g55127; PIDN:CAA26231.1; PID:g55128  
 C;Genetics:  
 A;Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3  
 C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
 C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F;1-20/Domain: signal sequence #status predicted <SIG>  
 F;21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
 F;32-63/Domain: EGF homology <EGF>  
 F;71-152/Domain: kringle homology <KRG>  
 F;180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
 F;180-421/Domain: trypsin homology <TRY>  
 F;169-301,211-227,219-290,315-384,347-363,374-402/disulfide bonds: #status predicted  
 F;226,277,378/Active site: His, Asp, Ser #status predicted  
 Query Match 75.9%; Score 387; DB 1; Length 433;  
 Best Local Similarity 72.7%; Pred. No. 1.2e-34;  
 Matches 64; Conservative 10; Mismatches 14; Indels 0; Gaps 0;  
 Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60  
 Db 69 KTCYHGNGDSYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 128  
 Qy 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88  
 Db 129 QRPCWYVQVGLRQFVQECMVHDCSLSK 156  
 RESULT 7  
 JS0599  
 t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat  
 N;Alternate names: tissue plasminogen activator  
 C;Species: Desmodus rotundus (common vampire bat)  
 C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
 C;Accession: JS0599  
 R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don  
 Gene 105, 229-237, 1991  
 A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des  
 A;Reference number: JS0597; MUID:92039036; PMID:1937019  
 A;Accession: JS0599  
 A;Molecule type: mRNA  
 A;Residues: 1-431 <KRA>  
 A;Cross-references: GB:M63989; NID:g166076; PIDN:AAA31594.1; PID:g166077  
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F;1-21/Domain: signal sequence #status predicted <SIG>  
 F;37-431/Product: propeptide #status predicted <PRO>  
 F;41-74/Domain: plasminogen activator beta #status predicted <PLA>  
 F;82-163/Domain: kringle homology <KRG>  
 F;180-425/Domain: trypsin homology <TRY>  
 F;41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide bon  
 F;135,352/binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;179-180/Cleavage site: His-Ser (plasmin) #status predicted  
 F;226,275,382/Active site: His, Asp, Ser #status predicted  
 F;345-361,378-406/Disulfide bonds: #status predicted  
 Query Match 47.3%; Score 241; DB 2; Length 431;  
 Best Local Similarity 50.0%; Pred. No. 1.1e-18;

Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;  
 Qy 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 61  
 Db 81 TCYKQDQVYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 140  
 Qy 62 RRPWCYVQVGLKPLVQECMVHDC 85  
 Db 141 SKPCWYVIKASKFILEFCSPVCS 164  
 RESULT 8  
 A34369  
 t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)  
 C;Species: Megaderma lyra  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C;Accession: A34369  
 R;Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jac  
 J. Biol. Chem. 264, 17947-17952, 1989  
 A;Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plas  
 A;Reference number: A34369; MUID:90036867; PMID:2509450  
 A;Accession: A34369  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-477 <GAR>  
 A;Cross-references: GB:J05082; NID:g166080; PIDN:AAA31596.1; PID:g166081  
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h  
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F;1-21/Domain: signal sequence #status predicted <SIG>  
 F;22-36/Domain: propeptide #status predicted <PRO>  
 F;37-477/Product: plasminogen activator #status predicted <PLA>  
 F;42-79/Domain: fibronectin type I repeat homology <1FA>  
 F;87-120/Domain: EGF homology <EGF>  
 F;128-209/Domain: kringle homology <KRG>  
 F;226-471/Domain: trypsin homology <TRY>  
 F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359  
 F;272,321,428/Active site: His, Asp, Ser #status predicted  
 Query Match 47.3%; Score 241; DB 1; Length 477;  
 Best Local Similarity 50.0%; Pred. No. 1.2e-18;  
 Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;  
 Qy 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 61  
 Db 127 TCYKQDQVYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 186  
 Qy 62 RRPWCYVQVGLKPLVQECMVHDC 85  
 Db 187 SKPCWYVIKASKFILEFCSPVCS 210  
 RESULT 9  
 JS0598  
 t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat  
 N;Alternate names: tissue plasminogen activator  
 C;Species: Desmodus rotundus (common vampire bat)  
 C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
 C;Accession: JS0598  
 R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; I  
 Gene 105, 229-237, 1991  
 A;Title: The plasminogen activator family from the salivary gland of the vampire bat I  
 A;Reference number: JS0597; MUID:92039036; PMID:1937019  
 A;Accession: JS0598  
 A;Molecule type: mRNA  
 A;Residues: 1-477 <KRA>  
 A;Cross-references: GB:M63988; NID:g166074; PIDN:AAA31593.1; PID:g166075  
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h  
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F;1-21/Domain: signal sequence #status predicted <SIG>  
 F;22-36/Domain: propeptide #status predicted <PRO>  
 F;37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>  
 F;42-79/Domain: fibronectin type I repeat homology <1FA>  
 F;87-120/Domain: EGF homology <EGF>



A:Molecule type: protein  
A:Residues: 33-45;311-320 <POH>  
A:Experimental source: uterus  
A:Note: In the uterus, cleavage of the activation peptide may also occur after 38-Gln  
R:Van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.  
J. Biol. Chem. 261, 14214-14218, 1986  
A:Reference number: A37567; MUID:87033611; PMID:3021732  
A:Contents: annotation; fibrin binding site  
R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger-  
EMBO J. 5, 3525-3530, 1986  
A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen ac-  
A:Reference number: A37568; MUID:87161761; PMID:3030730  
A:Contents: annotation; fibrin binding site  
R:Dodd, I.; Nunn, B.; Robinson, J.H.  
Thromb. Haemost. 59, 523-528, 1988  
A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type F  
A:Reference number: A60902; MUID:89044681; PMID:3142086  
A:Contents: annotation; novel forms of expressed recombinant t-PA  
R:Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opendakker, G.;  
Mol. Biol. Med. 3, 279-292, 1986  
A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expre-  
A:Reference number: A54645; MUID:86284200; PMID:3090401  
A:Accession: A54645  
A:Molecule type: mRNA  
A:Residues: 1-562 <BAR>  
A:Cross-references: GB:M15518; NID:G190031; PIDN:AAA60111.1; PID:G190032  
A:Note: parts of this sequence were confirmed by peptide sequencing  
R:Reddy, V.B.; Garramone, A.J.; Sasak, H.; Wei, C.  
DNA 6, 461-472, 1987  
A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells us-  
A:Reference number: I60110; MUID:88054470; PMID:2824147  
A:Accession: I60110  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-562 <RG>  
A:Cross-references: GB:M18182; NID:G340176; PIDN:AAA36800.1; PID:G340177  
R:Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.  
J. Biol. Chem. 260, 11223-11230, 1985  
A:Title: Isolation and characterization of the human tissue-type plasminogen activator S  
A:Reference number: I55232; MUID:95289338; PMID:3161893  
A:Accession: I55232  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-36 <RE>  
A:Cross-references: GB:M11890; NID:G339837; PIDN:AAA61213.1; PID:G339839  
C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single  
C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond.  
C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat.  
C:Genetics:  
A:Gene: GDB:PLAT  
A:Cross-references: GDB:I119496; OMIM:173370  
A:Map position: 8p12-8p12  
A:Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 510/1  
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homol  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-32/Domain: propeptide #status predicted <PRO>  
F:33-562/Product: t-plasminogen activator #status experimental <MAT>  
F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>  
F:41-78/Domain: fibronectin type I repeat homology <R1>  
F:86-119/Domain: EGF homology <EGF>  
F:127-208/Domain: kringle homology <KR1>  
F:125-296/Domain: kringle homology <KR2>  
F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>  
F:311-556/Domain: trypsin homology <TRY>  
F:41-71-69-78-86-97-91-108-110-119-127-208, 148-190, 179-203, 215-296, 236-278, 267-291, 299-4  
F:152,493/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:129/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:310-311/Cleavage site: Arg-Ile (plasmin, trypsin) (partial) #status experimental  
F:357,406/Active site: His, Asp #status predicted  
F:513/Active site: Ser #status experimental

Query Match

44.3%; Score 226; DB 1; Length 562;

C/Accession: JS0600  
R:Kraetzschmar, J.: Haendler, B.: Boidol, W.: Bringmann, P.: Alagon, A.: E

Best Local Similarity 47.7%; Pred.No. 6.2e-17;  
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

Qy      2 TCYEGNGHFRGKASTDTMGRPLCPNWSATVLQOITYHAHRSDALQLGLGHKNYC RNP DNR 61  
         |||||        |||||  
  
Db     126 TCYEDQGISVGTWSTAESAECTNNSSALAQPYSRGRRPD A IRLGLGNHN YC RN PDRD 185

Qy          62 RRWCYYQQVG L K PL V QEC MW HD CAGD 87  
         :||||:  
  
Db        186 SKWCYVFKA GYSSEFCSPACSEG 211  
         :||||:

## RESULT 12

A35029  
t-plasminogen activator (EC 3.4.21.68) precursor - rat  
C:Species: *Rattus norvegicus* (Norway rat)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A35029; A31597  
R:Feng, P.; Ohlsson, M.; Ny, T.  
J. Biol. Chem. 265, 2022-2027, 1990  
A:Title: The structure of the TATA-less rat tissue-type plasminogen activator gene. Sp  
A:Reference number: A35029; MUID:90130448; PMID:2105315  
A:Accession: A35029  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-559 <FEN>  
A:Cross-references: GB:M31197; NID:g207429; PIDN:AAA42261.1; PID:g207431; GB:J05226  
R:NY, T.; Leonardsson, G.; Hsueh, A.J.W.  
DNA 7, 671-677, 1988  
A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activa  
A:Reference number: A31597; MUID:89170114; PMID:3148445  
A:Accession: A31597  
A:Molecule type: mRNA  
A:Residues: 1-379, 'K', 381-559 <NVT>  
A:Cross-references: GB:M33697; NID:g530159; PIDN:AAA1812.1; PID:g530150  
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h  
C:Keywords: fibrinolysis; glycoprotein; hyaluronase; kringle; serine proteinase  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-29/Domain: propeptide #status predicted <PRO>  
F:30-559/Product: t-plasminogen activator #status predicted <MAT>  
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>  
F:38-75/Domain: fibronectin type I repeat homology <f1>  
F:83-116/Domain: EGF homology <EGF>  
F:124-205/Domain: kringle homology <KR1>  
F:213-294/Domain: kringle homology <KR2>  
F:303-559/Product: t-plasminogen activator chain B #status predicted <BCH>  
F:303-553/Domain: trypsin homology <TR>  
F:338-68,66-75,93-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297  
F:149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted  
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match	44.1%	Score 225;	DB 1;	Length 559;
Best Local Similarity	47.7%	Pred. No. 7.9e-17;		
Matches 41;	Conservative 10;	Mismatches 35;	Indels 0;	Gaps 0;
QY	2	TCYEGNHGFFRGKASTDTMGRCPLPWNSATVLTQOTYHAHRS DALQLGLGKHNYCRNPDNR	61	
DB	123	TCPEQGITYRGTWSTAENGAEICINWNSALLSQPYSA RPNATKLGIGHNYCRNPDNR	182	
QY	62	RRPCYIVQVGLKPLVQECMWHDCADG	87	
DB	183	VKPWCYVFKAGKYTFECSTPACPKG	208	

RESULT 13  
JS0600

t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat  
N:Alternate names: tissue plasminogen activator  
C:Species: *Desmodus rotundus* (common vampire bat)  
C:date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C:Accession: J50600  
R:Kraetzschmar, J., Haendler, B., Langer, G., Boidol, W., Brinmann, P., Al-

Gene 105, 229-237, 1991  
A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des  
A;Reference number: JS0597; MUID:92039036; PMID:1937019  
A;Accession: JS0600  
A;Molecule type: mRNA  
A;Residues: 1-394 <KRA>  
A;Cross-references: GB:M63990; NID:gl66078; PIDN:AAA1595.1; PID:gl66079  
A;Note: the authors translated the codon ATC for residue 75 as Thr  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-36/Domain: propeptide #status predicted <PRO>  
F;37-394/Product: plasminogen activator gamma #status predicted <PLA>  
F;45-126/Domain: kringle homology <KRG>  
F;143-388/Domain: trypsin homology <TRY>  
F;45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:  
F;142-143/Cleavage site: His-Ser (plasmin) #status predicted  
F;189,238,345/Active site: His, Asp, Ser #status predicted  
F;315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 43.1%; Score 220; DB 2; Length 394;  
Best Local Similarity 45.2%; Pred. No. 2e-16;  
Matches 38; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASDTWGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNP DNR 61  
Db 44 TCYKDGVTYRGWTSSTSGAQCINWNSLLIRRTYNGRMP EAVKLGHNHNYCRNP DGA 103

QY 62 RRPWCYVQVGLKPLVQECMVHDC A 85  
Db 104 SRPWCYVIRAKRTSBSCSVPVCS 127

RESULT 14  
A29941  
t-plasminogen activator (EC 3.4.21.68) precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A29941; S48205; S48207; S48206  
R;Ricklees, R.J.; Darrow, A.L.; Strickland, S.  
J. Biol. Chem. 263, 1563-1569, 1988  
A;Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mR  
A;Reference number: A29941; MUID:88087303; PMID:2826484  
A;Accession: A29941  
A;Molecule type: mRNA  
A;Residues: 1-559 <RIC>  
A;Cross-references: GB:J03520; NID:G202109; PIDN:AAA0470.1; PID:G202110  
R;Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.  
Eur. J. Biochem. 224, 863-871, 1994  
A;Title: Characterization of the murine plasma fibrinolytic system.  
A;Reference number: S48202; MUID:95010076; PMID:7523120  
A;Accession: S48205  
A;Molecule type: protein  
A;Residues: 33-37, 'X', 39-40 <LIJ>  
A;Accession: S48207  
A;Molecule type: protein  
A;Residues: 309-316 <LI2>  
A;Accession: S48206  
A;Molecule type: protein  
A;Residues: 33-37, 'X', 39-40 <LIW>  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-17/Domain: signal sequence #status predicted <SIG>  
F;18-29/Domain: propeptide #status predicted <PRO>  
F;30-559/Product: t-plasminogen activator #status predicted <MAT>  
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>  
F;38-75/Domain: fibronectin type I repeat homology <1F1>  
F;83-116/Domain: EGF homology <EGF>  
F;124-205/Domain: kringle homology <KR1>  
F;213-294/Domain: kringle homology <KR2>  
F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>  
A;Accession: JS0597  
A;Molecule type: mRNA  
A;Residues: 1-558 <HAS>  
C;Comment: This protein acts as serine protease.

F;149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted  
F;355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 42.9%; Score 219; DB 1; Length 559;  
Best Local Similarity 46.0%; Pred. No. 3.6e-16;  
Matches 40; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASDTWGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNP DNR 61  
Db 123 TCPEQGQTYRGWTSSTSGAQCINWNSVLSLKPYNARRPNAIKLGLGHNHNYCRNP DNR 182

QY 62 RRPWCYVQVGLKPLVQECMVHDC A 88  
Db 183 LRPWCYVFRAGKYTFEFCSTPACPKG K 209

RESULT 15  
JS0597  
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat  
N;Alternate names: tissue plasminogen activator  
C;Species: Desmodus rotundus (common vampire bat)  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C;Accession: JS0597  
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boigol, W.; Bringmann, P.; Alagon, A.; D  
Gene 105, 229-237, 1991  
A;Title: the plasminogen activator family from the salivary gland of the vampire bat D  
A;Reference number: JS0597; MUID:92039036; PMID:1937019  
A;Accession: JS0597  
A;Molecule type: mRNA  
A;Residues: 1-477 <KRA>  
A;Cross-references: GB:M63987; NID:gl66070; PIDN:AAA1591.1; PID:gl66071  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-36/Domain: propeptide #status predicted <PRO>  
F;37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>  
F;42-79/Domain: fibronectin type I repeat homology <1FA>  
F;87-120/Domain: EGF homology <EGF>  
F;128-209/Domain: kringle homology <KRG>  
F;226-471/Domain: trypsin homology <TRY>  
F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359  
F;153,398/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;225-226/Cleavage site: His-Ser (plasmin) #status predicted  
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 41.8%; Score 213; DB 2; Length 477;  
Best Local Similarity 46.4%; Pred. No. 1.4e-15;  
Matches 39; Conservative 10; Mismatches 35; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASDTWGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNP DNR 61  
Db 127 TCYEGQGVTYRGWTSSTSGAQCINWNSLLIRRTYNGRMP DAFNLGLGHNHNYCRNP DGA 186

QY 62 RRPWCYVQVGLKPLVQECMVHDC A 85  
Db 187 RRPWCYVIRAGKYTFEFCSTPACPKG K 210

RESULT 16  
JCS878  
plasma hyaluronan-binding protein precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 11-Mar-1998 #sequence\_revision 11-Mar-1998 #text\_change 16-Jul-1999  
C;Accession: JCS878  
R;Hashimoto, K.; Tobe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, I  
Biol. Pharm. Bull. 20, 1127-1130, 1997  
A;Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-b  
A;Reference number: JCS878; MUID:98065239; PMID:9401717  
A;Accession: JCS878  
A;Molecule type: mRNA  
A;Residues: 1-558 <HAS>  
C;Comment: This protein acts as serine protease.



```

F;373-407/Product: hepatocyte growth factor activator light chain #status experimental <
F;408-655/Product: hepatocyte growth factor activator heavy chain #status experimental <
F;408-641/Domain: trypsin homology <TRY>
F;40,48,290,468,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;164-175,169,186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-366
F;147,497,598/Active site: His, Asp, Ser #status predicted

Query Match 37.8%; Score 193; DB 1; Length 655;
Best Local Similarity 54.5%; Pred. No. 2.9e-13;
Matches 36; Conservative 4; Mismatches 26; Indels 0; Gaps 0;

Qy 3 CYEGNGHFYRGKASTDTMGSPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDRR 62
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 286 CFLGNGTGYRGVASTSAGLSCLANNSDLLYQELHVDVSGAALLGLGPHAYCRNPDNDE 345

Qy 63 RPWCYV 68
      :|||||
Db 346 RPWCYV 351

RESULT 20
S28941
coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)
N;Alternate names: Hageman factor
C;Species: Cavia porcellus (guinea pig)
C;Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000
C;Accession: S28941
R;Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.; Okabe, H.;
Biochim. Biophys. Acta 1159, 113-121, 1992
A;Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage site
A;Reference number: S28941; MUID:93003367; PMID:1390917
A;Accession: S28941
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-603 <SW>
A;Cross-references: EMBL:X68615; NID:g49578; PIDN:CAA48600.1; PID:g49579
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C;Keywords: hydrolase; serine proteinase
F;46-87/Domain: fibronectin type II repeat homology <1F2>
F;134-169/Domain: fibronectin type I repeat homology <FB1>
F;177-208/Domain: EGF homology <EGF>
F;216-294/Domain: kringle homology <KRG>
F;359-597/Domain: trypsin homology <TRY>

Query Match 35.4%; Score 180.5; DB 2; Length 603;
Best Local Similarity 42.9%; Pred. No. 6.2e-12;
Matches 36; Conservative 11; Mismatches 32; Indels 5; Gaps 2;

Qy 2 TCYEGNGHFYRGKASTDTMGSPCLPWNSATVLQOTYHAHRSD-ALQLGLGKHNYCRNPDRN 60
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 215 SCYEGRGVSYRGMAITTVSGAKCORNAS----EATYRNMTAEQALRGLGHHTFCRNPDRN 270

Qy 61 RRREWCYVQGLKPLVQECMVHDC 84
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 271 DTRPWCYVFMGNRLSWEYCDLAQC 294

RESULT 21
KFHU12
coagulation factor XIIa (EC 3.4.21.38) precursor [validated] - human
N;Alternate names: Hageman factor (activated)
C;Species: Homo sapiens (man)
C;Date: 27-Nov-1985 #sequence_revision 30-Jun-1991 #text_change 08-Dec-2000
C;Accession: A29411; A26814; A00930; A25191; A22248; A21037
R;Cool, D.E.; MacGillivray, R.T.A.
J. Biol. Chem. 262, 13662-13673, 1987
A;Title: Characterization of the human blood coagulation factor XII gene. Intron/exon gene
A;Reference number: A29411; MUID:88007593; PMID:2888762
A;Accession: A29411
A;Molecule type: DNA
A;Residues: 1-615 <CO>
A;Cross-references: GB:M17466; GB:J02807; NID:gl80355; PIDN:AAB59490.1; PID:gl80357
R;Tripod, M.; Citarella, F.; Guida, S.; Galeffi, P.; Fantoni, A.; Cortese, R.

```



		Matches	33;	Conservative	6;	Mismatches	24;	Indels	5;	Gaps	2;
QY	2	TCYEGNGHYFGKASTDTMGRPCLPWN	SATVQLQTY	HAHRSDALQ	LGKHN	CNP	DPN	60			
Db	216	SCYDGRGLSYGLARTLLSGAF	CQPWAS	----	EATYRN	TAEQARN	WGLG	HAF	CR	FP	DN
QY	61	RRRPWCYV	68								
Db	272	DIRPWCYV	279								

RESULT 22  
A32869  
apolipoprotein(a) (EC 3.4.21.-) - rhesus macaque (fragment)  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 22-Nov-1989 #sequence\_revision 22-Nov-1989 #text\_change 22-Jun-1999  
R:Tomlinson, J.B.; McLean, J.W.; Lawn, R.M.  
J. Biol. Chem. 264, 5957-5965, 1989  
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.  
A:Reference number: A32869; MUID:89174660; PMID:2925643  
A:Accession: A32869  
A:Molecule type: mRNA  
A:Residues: 1-1420 <TOM>  
A:Cross-references: GB:J04635; NID:9342072; PIDN:AAA36833.1; PID:G342073  
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology  
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase  
F:50-127/Domain: kringle homology <KR1>  
F:164-241/Domain: kringle homology <KR2>  
F:278-355/Domain: kringle homology <KR3>  
F:392-469/Domain: kringle homology <KR4>  
F:506-583/Domain: kringle homology <KR5>  
F:620-697/Domain: kringle homology <KR6>  
F:726-803/Domain: kringle homology <KR7>  
F:840-917/Domain: kringle homology <KR8>  
F:954-1031/Domain: kringle homology <KR9>  
F:1068-1145/Domain: kringle homology <KR10>  
F:1191-1413/Domain: trypsin homology <TRY>

Query Match 31.6%; Score 161; DB 2; Length 1420;  
Best Local Similarity 40.9%; Pred. No. 1.9e-09;  
Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;  
QY 3 CYEGNGHYFGKASTDTMGRPCLPWN SATVQLQTY HAHRSDALQ LGKHN CNP DN 60  
Db 1068 CYHGNGQSYRGTSITVTGTCQSWSSMT PHQHKRTPENHPNDLTM-----NYCRNPDA 1122  
QY 61 RRRPWCYVQVGLKPLVQ--CMVHDCAD 86  
Db 1123 DTGFWCFV---MDPSVRREYCNLTRCS 1147

RESULT 23  
E61545  
plasmin (EC 3.4.21.7) precursor - dog (fragments)  
N:Alternate names: plasminogen  
C:Species: Canis lupus familiaris (dog)  
C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 12-May-1995  
R:Schaller, J.; Rickli, E.E.  
Enzyme 40, 63-69, 1988  
A:Title: Structural aspects of the plasminogen of various species.  
A:Reference number: A61545; MUID:89005015; PMID:3168975  
A:Accession: E61545  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-120 <SCH>  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C:Keywords: hydrolase; kringle homology <KR1>  
F:37-114/Domain: kringle homology <KR1>

Query Match 30.6%; Score 156; DB 2; Length 120;  
Best Local Similarity 39.6%; Pred. No. 6.5e-10;  
Matches 36; Conservative 7; Mismatches 30; Indels 18; Gaps 5;  
QY 3 CYEGNGHYFGKASTDTMGRPCLPWN SATVQLQTY HAHRSDALQ LGKHN CNP DN 57  
Db 37 CYHGNGQSYRGTSITVTGTCQSWSSMT-----PHRHEKTPHFPEAGL-TWNYCRN 88  
QY 58 PDNRPPWCYVQVGLKPLV--QECMVHDCAD 86  
Db 89 PDADKSPWCYT---TDPVSRWFECNLRKCLD 116

RESULT 24  
S00657  
apolipoprotein(a) (EC 3.4.21.-) precursor [validated] - human  
N:Alternate names: apolipoprotein(a); lipoprotein(a) chain apo(a)  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 08-Dec-2000  
C:Accession: S00657; A28017; A47277; I60906; A47233; I52415; I65286  
R:McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.; Chen, E.Y.; Fless, G.M.; Sca  
Nature 330, 132-137, 1987  
A:Title: cDNA sequence of human apolipoprotein(a) is homologous to plasminogen.  
A:Reference number: S00657; MUID:88039109; PMID:3670400  
A:Accession: S00657  
A:Molecule type: mRNA  
A:Residues: 1-4548 <MCL>  
A:Cross-references: GB:X06290; EMBL:X06696; NID:928619; PIDN:CAA29618.1; PID:G28620  
R:Eaton, D.L.; Fless, G.M.; Kohr, W.J.; McLean, J.W.; Xu, Q.T.; Miller, C.G.; Lawn, R.M  
Proc. Natl. Acad. Sci. U.S.A. 84, 3224-3228, 1987  
A:Title: Partial amino acid sequence of apolipoprotein(a) shows that it is homologous t  
A:Reference number: A28017; MUID:87204109; PMID:3472206  
A:Accession: A28017  
A:Molecule type: protein  
A:Residues: 20-21, P', 23-34; 177-179, N', 181-186, T', 188-196, 'DKG', 200-292-314, W', 316-3  
X', 4396-4401 <EAT>  
R:Wade, D.P.; Clarke, J.G.; Lindahl, G.E.; Liu, A.C.; Zysow, B.R.; Meer, K.; Schwartz,  
Proc. Natl. Acad. Sci. U.S.A. 90, 1369-1373, 1993  
A:Title: 5' control regions of the apolipoprotein(a) gene and members of the related pl  
A:Reference number: A47277; MUID:93165698; PMID:7679504  
A:Accession: A47277  
A:Status: preliminary; translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-16 <RES>  
A:Cross-references: GB:I07899; NID:967973; PID:G967974  
R:Malgaretti, N.; Acquati, F.; Magnaghi, P.; Bruno, L.; Pontoglio, M.; Rocchi, M.; Sacc  
Proc. Natl. Acad. Sci. U.S.A. 89, 11584-11588, 1992  
A:Title: Characterization by yeast artificial chromosome cloning of the linked apolipof  
A:Reference number: A47233; MUID:93087573; PMID:1454851  
A:Accession: I60906  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-16 <RE2>  
A:Cross-references: GB:M90078; NID:9178786; PIDN:AAA35547.1; PID:G553188  
A:Note: apo(a) gene 1 (nomenclature of reference I52415)  
A:Accession: A47233  
A:Status: preliminary; translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-16 <RE3>  
A:Cross-references: GB:M90079; NID:9178784; PIDN:AAA35546.1; PID:G553187  
R:Ichinose, A.  
Biochemistry 31, 3113-3118, 1992  
A:Title: Multiple members of the plasminogen-apolipoprotein(a) gene family associated v  
A:Reference number: I52415; MUID:92207924; PMID:1554698  
A:Accession: I52415  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-16 <RE3>  
A:Cross-references: GB:M86877; NID:9178780; PIDN:AA849909.1; PID:G553185  
A:Note: apo(a) gene 1 (nomenclature of reference I52415)  
A:Accession: I65286  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-16 <RE4>  
A:Cross-references: GB:M86878; NID:9178782; PIDN:AAA51749.1; PID:G553186





N:Alternate names: plasminogen  
C:Species: Capra aegagrus hircus (domestic goat)  
C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 12-May-1995  
C:Accession: C61545  
R:Schaller, J.; Rickli, E.E.  
Enzyme 40, 63-69, 1988  
A:Title: Structural aspects of the plasminogen of various species.  
A:Reference number: A61545; MUID:89005015; PMID:3168975  
A:Accession: C61545  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-123 <SCH>  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C:Keywords: hydrolase; serine proteinase  
F:41-118/Domain: kringle homology <KR4>

Query Match 29.4%; Score 150; DB 2; Length 123;  
Best Local Similarity 37.5%; Pred. No. 3e-09;  
Matches 33; Conservative 12; Mismatches 31; Indels 12; Gaps 4;

QY 3 CYEGNGHFRGKASTDTMGRCPLPWNATV--LQOTYHAHRSDALQLGLGKHNCRNPD 60  
DB 41 CYHNGQSYRGTSSTVTGRKQOSWSSMTPHRHQKTPESYPNAGLTM-----NYCRNPD 95  
QY 61 RRRPWCYVQVGLKPLV--QECMVHDCAD 86  
DB 96 DKSPWCYT---TDPRVWEFCNLKCKSE 120

RESULT 28  
A40522  
plasmin (EC 3.4.21.7) precursor - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 28-Feb-1992 #sequence\_revision 17-Apr-1993 #text\_change 16-Jul-1999  
C:Accession: A40522  
R:Kanagas, J.J.; Makker, S.P.  
J. Biol. Chem. 266, 10825-10829, 1991  
A:Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor  
A:Reference number: A40522; MUID:91250378; PMID:1645711  
A:Accession: A40522  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-169 <KAN>  
A:Cross-references: GB:M62832; NID:G206215; PIDN:AAA41884.1; PID:G554488  
A:Note: the authors translated the codon TCT for residue 76 as Ala  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:34-112/Domain: kringle homology <KRG>  
F:34-112/55-95,83-107/Disulfide bonds: #status predicted

Query Match 29.3%; Score 149.5; DB 2; Length 169;  
Best Local Similarity 37.8%; Pred. No. 4.6e-09;  
Matches 34; Conservative 12; Mismatches 29; Indels 15; Gaps 6;

QY 3 CYEGNGHFRGKASTDTMGRCPLPWNATV--LQOTYHAHRSDALQL---GLGKHNCRNPD 59  
DB 34 CYQNGKSYRGTSSTNTGKQOSW-----VSMTPHSHSKTPANFPDSSL-EMNYCRNPD 87  
QY 60 NRRPWCYVQVGLKPLV--QECMVHDCAD 86  
DB 88 NDORGWCFT---TDFSRWEYCNLKKCKSE 114

RESULT 29  
PLBO  
plasmin (EC 3.4.21.7) precursor - bovine  
N:Alternate names: plasminogen  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Sep-1987 #sequence\_revision 28-Apr-1995 #text\_change 18-Jun-1999  
C:Accession: S45046; A25835; I45961; S03736  
R:Berglund, L.; Andersen, M.D.; Petersen, T.E.  
submitted to the EMBL Data Library, May 1994  
A:Description: Cloning and characterization of the bovine plasminogen cDNA.

A:Reference number: S45046  
A:Accession: S45046  
A:Molecule type: mRNA  
A:Residues: 1-812 <BER>  
A:Cross-references: EMBL:X79402; NID:G494962; PIDN:CAA55939.1; PID:G494963  
A:Experimental source: liver  
A:Note: it is uncertain whether Met-1 or Met-8 is the initiator  
R:Schaller, J.; Moser, P.W.; Danneberger-Muller, G.A.K.; Rosselet, S.J.; Kampfer, U.; Rickli, J. Biochem. 149, 267-278, 1985  
A:Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plasminogen  
A:Reference number: A25835; MUID:85203906; PMID:3846532  
A:Accession: A25835  
A:Molecule type: protein  
A:Residues: 27-334, 'D', '336-515, 'H', '517-554, 'L', '556-812 <SCH>  
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.  
Biochemistry 23, 4243-4250, 1984  
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and bovine plasminogen  
A:Reference number: I45961; MUID:85023311; PMID:6148961  
A:Accession: I45961  
A:Status: translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 706-743, 'R', '745-812 <MAL>  
R:Cross-references: GB:K02935; NID:G163551; PIDN:AAA30714.1; PID:G163552  
R:Brumsholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, J.E.; Lergier, W.; Manneberg, M.  
Eur. J. Biochem. 114, 465-470, 1981  
A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human, bovine and rat plasminogen  
A:Reference number: S03736; MUID:81212097; PMID:7238497  
A:Accession: S03736  
A:Molecule type: protein  
A:Residues: 27-83 <BRU>  
C:Function:

A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of the walls of the graafian follicle; also activates the urokinase-type plasminogen activator  
A:Pathway: fibrinolysis  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C:Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma  
F:8-103/Domain: signal sequence #status predicted <SIG>  
F:1-26/Domain: plasminogen-related protein precursor homology <PRO>  
F:27-812/Product: plasminogen #status experimental <APT>  
F:27-103/Domain: activation peptide #status experimental <APT>  
F:104-583,584-812/Product: plasmin #status experimental <MAT>  
F:104-583/Domain: plasmin chain A #status experimental <ACH>  
F:110-188/Domain: kringle homology <KR1>  
F:192-269/Domain: kringle homology <KR2>  
F:282-359/Domain: kringle homology <KR3>  
F:384-461/Domain: kringle homology <KR4>  
F:485-564/Domain: kringle homology <KR5>  
F:584-812/Domain: plasmin chain B #status experimental <BCH>  
F:584-805/Domain: trypsin homology <TRY>  
F:56-80,60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,303-342,343-344,345-346,347-348,349-350,351-352,353-354,355-356,357-358,359-360,361-362,363-364,365-366,367-368,369-370,371-372,373-374,375-376,377-378,379-380,381-382,383-384,385-386,387-388,389-390,391-392,393-394,395-396,397-398,399-400,401-402,403-404,405-406,407-408,409-410,411-412,413-414,415-416,417-418,419-420,421-422,423-424,425-426,427-428,429-430,431-432,433-434,435-436,437-438,439-440,441-442,443-444,445-446,447-448,449-450,451-452,453-454,455-456,457-458,459-460,461-462,463-464,465-466,467-468,469-470,471-472,473-474,475-476,477-478,479-480,481-482,483-484,485-486,487-488,489-490,491-492,493-494,495-496,497-498,499-500,501-502,503-504,505-506,507-508,509-510,511-512,513-514,515-516,517-518,519-520,521-522,523-524,525-526,527-528,529-530,531-532,533-534,535-536,537-538,539-540,541-542,543-544,545-546,547-548,549-550,551-552,553-554,555-556,557-558,559-560,561-562,563-564,565-566,567-568,569-570,571-572,573-574,575-576,577-578,579-580,581-582,583-584,585-586,587-588,589-590,591-592,593-594,595-596,597-598,599-600,601-602,603-604,605-606,607-608,609-610,611-612,613-614,615-616,617-618,619-620,621-622,623-624,625-626,627-628,629-630,631-632,633-634,635-636,637-638,639-640,641-642,643-644,645-646,647-648,649-650,651-652,653-654,655-656,657-658,659-660,661-662,663-664,665-666,667-668,669-670,671-672,673-674,675-676,677-678,679-680,681-682,683-684,685-686,687-688,689-690,691-692,693-694,695-696,697-698,699-700,701-702,703-704,705-706,707-708,709-710,711-712,713-714,715-716,717-718,719-720,721-722,723-724,725-726,727-728,729-730,731-732,733-734,735-736,737-738,739-740,741-742,743-744,745-746,747-748,749-750,751-752,753-754,755-756,757-758,759-760,761-762,763-764,765-766,767-768,769-770,771-772,773-774,775-776,777-778,779-780,781-782,783-784,785-786,787-788,789-790,791-792,793-794,795-796,797-798,799-800,801-802,803-804,805-806,807-808,809-810,811-812,813-814,815-816,817-818,819-820,821-822,823-824,825-826,827-828,829-830,831-832,833-834,835-836,837-838,839-840,841-842,843-844,845-846,847-848,849-850,851-852,853-854,855-856,857-858,859-860,861-862,863-864,865-866,867-868,869-870,871-872,873-874,875-876,877-878,879-880,881-882,883-884,885-886,887-888,889-890,891-892,893-894,895-896,897-898,899-900,901-902,903-904,905-906,907-908,909-910,911-912,913-914,915-916,917-918,919-920,921-922,923-924,925-926,927-928,929-930,931-932,933-934,935-936,937-938,939-940,941-942,943-944,945-946,947-948,949-950,951-952,953-954,955-956,957-958,959-960,961-962,963-964,965-966,967-968,969-970,971-972,973-974,975-976,977-978,979-980,981-982,983-984,985-986,987-988,989-990,991-992,993-994,995-996,997-998,999-1000,1001-1002,1003-1004,1005-1006,1007-1008,1009-1010,1011-1012,1013-1014,1015-1016,1017-1018,1019-1020,1021-1022,1023-1024,1025-1026,1027-1028,1029-1030,1031-1032,1033-1034,1035-1036,1037-1038,1039-1040,1041-1042,1043-1044,1045-1046,1047-1048,1049-1050,1051-1052,1053-1054,1055-1056,1057-1058,1059-1060,1061-1062,1063-1064,1065-1066,1067-1068,1069-1070,1071-1072,1073-1074,1075-1076,1077-1078,1079-1080,1081-1082,1083-1084,1085-1086,1087-1088,1089-1090,1091-1092,1093-1094,1095-1096,1097-1098,1099-1100,1101-1102,1103-1104,1105-1106,1107-1108,1109-1110,1111-1112,1113-1114,1115-1116,1117-1118,1119-1120,1121-1122,1123-1124,1125-1126,1127-1128,1129-1130,1131-1132,1133-1134,1135-1136,1137-1138,1139-1140,1141-1142,1143-1144,1145-1146,1147-1148,1149-1150,1151-1152,1153-1154,1155-1156,1157-1158,1159-1160,1161-1162,1163-1164,1165-1166,1167-1168,1169-1170,1171-1172,1173-1174,1175-1176,1177-1178,1179-1180,1181-1182,1183-1184,1185-1186,1187-1188,1189-1190,1191-1192,1193-1194,1195-1196,1197-1198,1199-1200,1201-1202,1203-1204,1205-1206,1207-1208,1209-1210,1211-1212,1213-1214,1215-1216,1217-1218,1219-1220,1221-1222,1223-1224,1225-1226,1227-1228,1229-1230,1231-1232,1233-1234,1235-1236,1237-1238,1239-1240,1241-1242,1243-1244,1245-1246,1247-1248,1249-1250,1251-1252,1253-1254,1255-1256,1257-1258,1259-1260,1261-1262,1263-1264,1265-1266,1267-1268,1269-1270,1271-1272,1273-1274,1275-1276,1277-1278,1279-1280,1281-1282,1283-1284,1285-1286,1287-1288,1289-1290,1291-1292,1293-1294,1295-1296,1297-1298,1299-1300,1301-1302,1303-1304,1305-1306,1307-1308,1309-1310,1311-1312,1313-1314,1315-1316,1317-1318,1319-1320,1321-1322,1323-1324,1325-1326,1327-1328,1329-1330,1331-1332,1333-1334,1335-1336,1337-1338,1339-1340,1341-1342,1343-1344,1345-1346,1347-1348,1349-1350,1351-1352,1353-1354,1355-1356,1357-1358,1359-1360,1361-1362,1363-1364,1365-1366,1367-1368,1369-1370,1371-1372,1373-1374,1375-1376,1377-1378,1379-1380,1381-1382,1383-1384,1385-1386,1387-1388,1389-1390,1391-1392,1393-1394,1395-1396,1397-1398,1399-1400,1401-1402,1403-1404,1405-1406,1407-1408,1409-1410,1411-1412,1413-1414,1415-1416,1417-1418,1419-1420,1421-1422,1423-1424,1425-1426,1427-1428,1429-1430,1431-1432,1433-1434,1435-1436,1437-1438,1439-1440,1441-1442,1443-1444,1445-1446,1447-1448,1449-1450,1451-1452,1453-1454,1455-1456,1457-1458,1459-1460,1461-1462,1463-1464,1465-1466,1467-1468,1469-1470,1471-1472,1473-1474,1475-1476,1477-1478,1479-1480,1481-1482,1483-1484,1485-1486,1487-1488,1489-1490,1491-1492,1493-1494,1495-1496,1497-1498,1499-1500,1501-1502,1503-1504,1505-1506,1507-1508,1509-1510,1511-1512,1513-1514,1515-1516,1517-1518,1519-1520,1521-1522,1523-1524,1525-1526,1527-1528,1529-1530,1531-1532,1533-1534,1535-1536,1537-1538,1539-1540,1541-1542,1543-1544,1545-1546,1547-1548,1549-1550,1551-1552,1553-1554,1555-1556,1557-1558,1559-1560,1561-1562,1563-1564,1565-1566,1567-1568,1569-1570,1571-1572,1573-1574,1575-1576,1577-1578,1579-1580,1581-1582,1583-1584,1585-1586,1587-1588,1589-1590,1591-1592,1593-1594,1595-1596,1597-1598,1599-1600,1601-1602,1603-1604,1605-1606,1607-1608,1609-1610,1611-1612,1613-1614,1615-1616,1617-1618,1619-1620,1621-1622,1623-1624,1625-1626,1627-1628,1629-1630,1631-1632,1633-1634,1635-1636,1637-1638,1639-1640,1641-1642,1643-1644,1645-1646,1647-1648,1649-1650,1651-1652,1653-1654,1655-1656,1657-1658,1659-1660,1661-1662,1663-1664,1665-1666,1667-1668,1669-1670,1671-1672,1673-1674,1675-1676,1677-1678,1679-1680,1681-1682,1683-1684,1685-1686,1687-1688,1689-1690,1691-1692,1693-1694,1695-1696,1697-1698,1699-1700,1701-1702,1703-1704,1705-1706,1707-1708,1709-1710,1711-1712,1713-1714,1715-1716,1717-1718,1719-1720,1721-1722,1723-1724,1725-1726,1727-1728,1729-1730,1731-1732,1733-1734,1735-1736,1737-1738,1739-1740,1741-1742,1743-1744,1745-1746,1747-1748,1749-1750,1751-1752,1753-1754,1755-1756,1757-1758,1759-1760,1761-1762,1763-1764,1765-1766,1767-1768,1769-1770,1771-1772,1773-1774,1775-1776,1777-1778,1779-1780,1781-1782,1783-1784,1785-1786,1787-1788,1789-1790,1791-1792,1793-1794,1795-1796,1797-1798,1799-1800,1801-1802,1803-1804,1805-1806,1807-1808,1809-1810,1811-1812,1813-1814,1815-1816,1817-1818,1819-1820,1821-1822,1823-1824,1825-1826,1827-1828,1829-1830,1831-1832,1833-1834,1835-1836,1837-1838,1839-1840,1841-1842,1843-1844,1845-1846,1847-1848,1849-1850,1851-1852,1853-1854,1855-1856,1857-1858,1859-1860,1861-1862,1863-1864,1865-1866,1867-1868,1869-1870,1871-1872,1873-1874,1875-1876,1877-1878,1879-1880,1881-1882,1883-1884,1885-1886,1887-1888,1889-1890,1891-1892,1893-1894,1895-1896,1897-1898,1899-1900,1901-1902,1903-1904,1905-1906,1907-1908,1909-1910,1911-1912,1913-1914,1915-1916,1917-1918,1919-1920,1921-1922,1923-1924,1925-1926,1927-1928,1929-1930,1931-1932,1933-1934,1935-1936,1937-1938,1939-1940,1941-1942,1943-1944,1945-1946,1947-1948,1949-1950,1951-1952,1953-1954,1955-1956,1957-1958,1959-1960,1961-1962,1963-1964,1965-1966,1967-1968,1969-1970,1971-1972,1973-1974,1975-1976,1977-1978,1979-1980,1981-1982,1983-1984,1985-1986,1987-1988,1989-1990,1991-1992,1993-1994,1995-1996,1997-1998,1999-2000,2001-2002,2003-2004,2005-2006,2007-2008,2009-2010,2011-2012,2013-2014,2015-2016,2017-2018,2019-2020,2021-2022,2023-2024,2025-2026,2027-2028,2029-2030,2031-2032,2033-2034,2035-2036,2037-2038,2039-2040,2041-2042,2043-2044,2045-2046,2047-2048,2049-2050,2051-2052,2053-2054,2055-2056,2057-2058,2059-2060,2061-2062,2063-2064,2065-2066,2067-2068,2069-2070,2071-2072,2073-2074,2075-2076,2077-2078,2079-2080,2081-2082,2083-2084,2085-2086,2087-2088,2089-2090,2091-2092,2093-2094,2095-2096,2097-2098,2099-2100,2101-2102,2103-2104,2105-2106,2107-2108,2109-2110,2111-2112,2113-2114,2115-2116,2117-2118,2119-2120,2121-2122,2123-2124,2125-2126,2127-2128,2129-2130,2131-2132,2133-2134,2135-2136,2137-2138,2139-2140,2141-2142,2143-2144,2145-2146,2147-2148,2149-2150,2151-2152,2153-2154,2155-2156,2157-2158,2159-2160,2161-2162,2163-2164,2165-2166,2167-2168,2169-2170,2171-2172,2173-2174,2175-2176,2177-2178,2179-2180,2181-2182,2183-2184,2185-2186,2187-2188,2189-2190,2191-2192,2193-2194,2195-2196,2197-2198,2199-2200,2201-2202,2203-2204,2205-2206,2207-2208,2209-2210,2211-2212,2213-2214,2215-2216,2217-2218,2219-2220,2221-2222,2223-2224,2225-2226,2227-2228,2229-2230,2231-2232,2233-2234,2235-2236,2237-2238,2239-2240,2241-2242,2243-2244,2245-2246,2247-2248,2249-2250,2251-2252,2253-2254,2255-2256,2257-2258,2259-2260,2261-2262,2263-2264,2265-2266,2267-2268,2269-2270,2271-2272,2273-2274,2275-2276,2277-2278,2279-2280,2281-2282,2283-2284,2285-2286,2287-2288,2289-2290,2291-2292,2293-2294,2295-2296,2297-2298,2299-2300,2301-2302,2303-2304,2305-2306,2307-2308,2309-2310,2311-2312,2313-2314,2315-2316,2317-2318,2319-2320,2321-2322,2323-2324,2325-2326,2327-2328,2329-2330,2331-2332,2333-2334,2335-2336,2337-2338,2339-2340,2341-2342,2343-2344,2345-2346,2347-2348,2349-2350,2351-2352,2353-2354,2355-2356,2357-2358,2359-2360,2361-2362,2363-2364,2365-2366,2367-2368,

Search completed: December 3, 2003, 14:44:12  
Job time : 4.55172 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:53 ; Search time 2.82759 Seconds  
(without alignments)  
1463.563 Million cell updates/sec

Title: US-09-880-503-1

Perfect score: 510

Sequence: 1 KTCYEGNGHFYRGKASTDM.....QVGLKPLVQECMVHDCADGK 88

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	431	UROK_HUMAN	P00749 homo sapien
2	475	93.1	433	UROK_PAPY	P16277 papio cynoc
3	420	82.4	442	UROK_PIG	P04185 sus scrofa
4	406	79.6	432	UROK_RAT	P29598 rattus norv
5	396	77.6	433	UROK_BOVIN	P05589 bos taurus
6	387	75.9	433	UROK_MOUSE	P06869 mus musculu
7	241	47.3	431	URTE_DESRO	P98121 desmodus ro
8	226	44.3	562	TPA_HUMAN	P15638 desmodus ro
9	225	44.1	559	TPA_RAT	P00750 homo sapien
10	220	43.1	394	URTG_DESRO	P19637 rattus norv
11	219	42.9	559	TPA_MOUSE	P49150 desmodus ro
12	213	41.8	477	URT1_DESRO	P11214 mus musculu
13	209	41.0	566	TPA_BOVIN	P98119 desmodus ro
14	198.5	38.9	434	UROK_CHICK	P28198 bos taurus
15	193	37.8	653	HGFA_HUMAN	P15120 gallus gall
16	188	36.9	653	HGFA_MOUSE	P04756 homo sapien
17	180.5	35.4	603	FA12_CAVPO	P09098 mus musculu
18	170.5	33.4	615	FA12_HUMAN	P04962 cavia porce
19	161	31.6	1420	APOA_MACMU	P00748 homo sapien
20	159	31.2	473	KRM1_MOUSE	P14417 macaca mula
21	159	31.2	473	KRM1_RAT	P99N43 mus musculu
22	159	31.2	473	KRM1_XENLA	P92484 rattus norv
23	156	30.6	452	KRM1_HUMAN	P09090 xenopus lae
24	156	30.6	475	APOA_HUMAN	P96mu8 homo sapien
25	155.5	30.5	533	FA12_BOVIN	P08519 homo sapien
26	149.5	29.3	169	PLMN_RAT	P98140 bos taurus
27	149.5	29.3	169	PLMN_BOVIN	P01177 rattus norv
28	149	29.2	812	PLMN_BOVIN	P06868 bos taurus
29	148.5	29.1	462	KRM2_HUMAN	P8ncw0 homo sapien
30	148.5	29.1	810	PLMN_ERIEU	P029485 erinaceus e
31	148	29.0	790	PLMN_PIG	P06867 sus scrofa
32	147.5	28.9	943	ROR2_HUMAN	P01974 homo sapien
33	147.5	28.9	944	ROR2_MOUSE	P92138 mus musculu

#### RESULT 1

UROK_HUMAN	STANDARD;	PRT;	431 AA.
ID	UROK_HUMAN	Q16618; Q969W6;	
AC	P00749; Q16618; Q969W6;		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	20-MAR-1987 (Rel. 04, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)		
DE	(U-plasminogen activator).		
GN	PLAU		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=85215647; PubMed=2987867;		
RA	Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Biasi F.;		
RT	"The human urokinase-plasminogen activator gene and its promoter.";		
RL	Nucleic Acids Res. 13:2759-2771(1985).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,		
RA	Steffens G.J., Heyneker H.L.;		
RT	"Cloning and expression of the gene for pro-urokinase in Escherichia coli.";		
RL	Biotechnology 3:923-929(1985).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86056954; PubMed=2415429;		
RA	Nagai M., Hiranatsu R., Kaneda T., Hayasuke N., Arimura H.,		
RA	Nishida M., Suyama T.;		
RT	"Molecular cloning of cDNA coding for human prepro-urokinase.";		
RL	Gene 36:183-188(1985).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=85203359; PubMed=3888571;		
RA	Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,		
RA	van Elsen A., Herzog A., Bollen A.;		
RT	"Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA.";		
RL	DNA 4:139-146(1985).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22338257; PubMed=12477932;		
RA	Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,		
RA	Nickerson D.A.;		
RT	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.		
RL	TISSUE=Lung;		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22338257; PubMed=12477932;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Horg L.,		

34	146.5	28.7	461	1	KRM2_MOUSE
35	146	28.6	937	1	ROR1_HUMAN
36	146	28.6	937	1	ROR1_MOUSE
37	144	28.2	810	1	PLMN_HUMAN
38	142.5	27.9	711	1	HGFL_HUMAN
39	141	27.6	333	1	PLMN_CANFA
40	141	27.6	810	1	PLMN_MACMU
41	139	27.3	685	1	ROR1_DROME
42	137.5	27.0	728	1	HGF_MOUSE
43	135.5	26.6	728	1	HGF_RAT
44	133.5	26.2	728	1	HGF_HUMAN
45	132	25.9	716	1	HGFL_MOUSE

#### ALIGNMENTS

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carinci P., Prange C., Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Touchman J.W., Schmutz J., Myers R.M., Butterfield Y.S.N., Kraywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.F., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[7] SEQUENCE OF 66-431 FROM N.A. MEDLINE=84272706; PubMed=6589620; Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.; "Identification and primary sequence of an unspliced human urokinase poly(A)+ RNA."; Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).

[8] SEQUENCE OF 21-177. MEDLINE=83055084; PubMed=6754569; Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E., Flohe L.; "The primary structure of high molecular mass urokinase from human urine. The complete amino acid sequence of the A chain."; Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).

[9] SEQUENCE OF 156-176 AND 179-224. MEDLINE=83003608; PubMed=6749491; Schaller J., Nick H., Rickli E.E., Gillessen D., Iergier W., Studer R.O.; "Human low-molecular-weight urinary urokinase. Partial characterization and preliminary sequence data of the two polypeptide chains."; Eur. J. Biochem. 125:251-257(1982).

[10] SEQUENCE OF 158-410. MEDLINE=83055099; PubMed=6754572; Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.; "The complete amino acid sequence of low molecular mass urokinase from human urine."; Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).

[11] X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS). MEDLINE=96000858; PubMed=8591045; Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D., Dobson C.M., Stuart D.I., Jones E.Y.; "The crystal structure of the catalytic domain of human urokinase-type plasminogen activator."; Structure 3:681-691(1995).

[12] X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411. MEDLINE=20266327; PubMed=10805774; Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G., Bode W., Magdolen V., Huber R., Moroder L.; "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly selective inhibitors of human urokinase."; Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).

[13] STRUCTURE BY NMR. MEDLINE=89127526; PubMed=2536903; Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.; "Dynamics of the multidomain fibrinolytic protein urokinase from two-dimensional NMR."; Nature 337:579-582(1989).

[14] STRUCTURE BY NMR OF 67-155. MEDLINE=93003110; PubMed=13271118; Li X., Smith R.A.G., Dobson C.M.;

"Sequential 1H NMR assignments and secondary structure of the kringle domain from urokinase."; Biochemistry 31:9562-9571(1992).

[15] STRUCTURE BY NMR OF 67-155. MEDLINE=94149701; PubMed=8107091; Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.; "Solution structure of the kringle domain from urokinase-type plasminogen activator."; J. Mol. Biol. 235:1548-1559(1994).

[16] VARIANT LEU-141. MEDLINE=96186279; PubMed=8652631; Yoshimoto M., Ushiyama Y., Sakai M., Tanaki S., Hara H., Takahashi K., Sawasaki Y., Hanada K.; "Characterization of single chain urokinase-type plasminogen activator with a novel amino-acid substitution in the kringle structure."; Biochim. Biophys. Acta 1293:83-89(1996).

[17] VARIANT LEU-141. MEDLINE=97218551; PubMed=9065989; Conne B., Berczy M., Belin D.; "Detection of polymorphisms in the human urokinase-type plasminogen activator gene."; Thromb. Haemost. 77:434-435(1997).

[18] ERRATUM. Conne B., Berczy M., Belin D.; Thromb. Haemost. 78:973-973(1997).

[19] VARIANT LEU-141. MEDLINE=97337920; PubMed=9194591; Turkmen B., Schmitt M., Schmalfeidt B., Trommler P., Hell W., Creutzburg S., Graeff H., Magdolen V.; "Mutational analysis of the genes encoding urokinase-type plasminogen activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer."; Electrophoresis 18:686-689(1997).

CC -!- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR THERAPY OF THROMBOLYTIC DISORDERS.

CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.

CC -!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used in Pulmonary Embolism (PE) to initiate fibrinolysis.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: Contains 1 kringle domain.

CC -!- SIMILARITY: Contains 1 EGF-like domain.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; X02419; CAA26268.1; -

DR EMBL; M15476; AAA61253.1; -

DR EMBL; D00244; BAA00175.1; -

DR EMBL; D11143; BAA01919.1; -

DR EMBL; X02760; CAA26535.1; -

DR EMBL; AF377330; AAK53822.1; -

DR EMBL; BC013575; AAH13575.1; -

DR EMBL; K03226; AAC97138.1; -

DR EMBL; K02286; AAA61252.1; -

DR EMBL; A21571; CAA01559.1; -

DR EMBL; A18397; CAA01390.1; -

DR PIR; A00931; UKHU.

DR PDB; 1XDU; 31-OCT-93.

```
Query Match      100.0%; Score 510; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.6e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFRGKASTDTMGRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60
DB 68 KTCYEGNGHFRGKASTDTMGRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 2
UROK_PAPCY STANDARD; PRT; 433 AA.
ID UROK_PAPCY
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAÜ.
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RT plasminogen activator."
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X51935; CAA36200.1;
DB PIR; S14587; UKBAY.
DB HSP; P00749; ILMW.
DB MEROPS; S01.231;
DB InterPro; IPR001314; Chymotrypsin.
DB InterPro; IPR006209; EGF like.
DB InterPro; IPR006210; IEGF.
DB InterPro; IPR000001; Kringle.
DB InterPro; IPR001254; Ser. protease_Try.
DB Pfam; PF00051; Kringle; 1.
DB PRINTS; PR00722; CHYMOTRYPSIN.
DB PRINTS; PR00018; KRINGLE.
DB ProDom; PD000395; Kringle; 1.
DB SMART; SM00181; EGF; 1.
DB SMART; SM00130; KR; 1.
DB SMART; SM00020; Tryp_Spc; 1.
DB PROSITE; PS00022; EGF_1; 1.
```

```
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 26 62 EGF-LIKE.
FT DOMAIN 69 150 KRINGLE.
FT DOMAIN 151 177 CONNECTING PEPTIDE.
FT DOMAIN 178 433 SERINE PROTEASE.
FT DISULFID 30 38 BY SIMILARITY.
FT DISULFID 32 50 BY SIMILARITY.
FT DISULFID 52 61 BY SIMILARITY.
FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
FT DISULFID 208 224 BY SIMILARITY.
FT DISULFID 216 287 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match      93.1%; Score 475; DB 1; Length 433;
Best Local Similarity 94.3%; Pred. No. 1.8e-47;
Matches 83; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFRGKASTDTMGRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60
DB 67 KTCYEGNGHFRGKASTDTMGRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 126

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 127 RRRPWCYVQVGLKPLVQECMVHDCADGK 154

RESULT 3
UROK_PIG STANDARD; PRT; 442 AA.
ID UROK_PIG
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAÜ.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator."
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
```

```

CC CC -!- SIMILARITY: Contains 1 kringle domain.
CC CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; X01648; CAA25806.1; -.
CC DR EMBL; X02724; CAA26511.1; -.
CC DR PIR; A00932; UKPG.
CC DR HSSP; P00749; 1KDU.
CC DR MEROPS; S01.231; -.
CC DR InterPro; IPR001314; Chymotrypsin.
CC DR InterPro; IPR006209; EGF like.
CC DR InterPro; IPR000001; Kringle.
CC DR InterPro; IPR001254; Ser_protease_Try.
CC DR Pfam; PF00051; kringle; 1.
CC DR Pfam; PF00089; trypsin; 1.
CC DR PRINTS; PR00722; CHYMOTRYPSIN.
CC DR PRINTS; PR00018; KRINGLE.
CC DR ProDom; PD000395; Kringle; 1.
CC DR SMART; SM00130; KR; 1.
CC DR SMART; SM00020; Tryp_Spc; 1.
CC DR PROSITE; PS00022; EGF_1; 1.
CC DR PROSITE; PS01186; EGF_2; FALSE NEG.
CC DR PROSITE; PS00021; KRINGLE_1; 1.
CC DR PROSITE; PS50070; KRINGLE_2; 1.
CC DR PROSITE; PS50240; TRYPSIN_DOM; 1.
CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
CC KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
CC KW Kringle; EGF-like domain; Zymogen; Signal.
CC FT SIGNAL 1 20
CC FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
CC FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
CC FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
CC FT CHAIN 21 65 EGF-LIKE.
CC FT DOMAIN 29 65 EGF-LIKE.
CC FT DOMAIN 72 153 KRINGLE.
CC FT DOMAIN 154 189 CONNECTING PEPTIDE.
CC FT DOMAIN 190 442 SERINE PROTEASE.
CC FT CARBOHYD 152 152 N-LINKED (GLCNAc. . .).
CC FT DISULFID 33 41 BY SIMILARITY.
CC FT DISULFID 35 53 BY SIMILARITY.
CC FT DISULFID 55 64 INTERCHAIN (BY SIMILARITY).
CC FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
CC FT DISULFID 220 236 BY SIMILARITY.
CC FT DISULFID 228 299 BY SIMILARITY.
CC FT DISULFID 324 393 BY SIMILARITY.
CC FT DISULFID 356 372 BY SIMILARITY.
CC FT DISULFID 383 411 BY SIMILARITY.
CC FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
CC FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
CC FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
CC FT CONFLICT 241 241 Q -> H (IN REF. 1; CAA25806).
CC FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA25811).
CC FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
CC SQ SEQUENCE 442 AA; 49116 MW; EE32PCEFS01321EE CRC64;
Query Match 82.4%; Score 420; DB 1; Length 442;
Best Local Similarity 81.8%; Pred. No. 4,1e-41;
Matches 72; Conservative 8; Mismatches 8; Indels 0; Gaps 0;
QY 1 KTCVEGNHGYRGKASTDMGRCLPWNASATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
70 QTCFEGNHSYRGKANTTGRPCLPWNASATVLLNTYHAHSPDALQLGLGKHNYCRNPDN 129
QY 61 RRRPCYVQGLKPLVQECMVHDCADGK 88
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 QRRPCYVQVGLKQLVQECMVPCNSGGE 157

```

## RESULT 4

```

UROK_RAT
ID UROK_RAT STANDARD; PRT; 432 AA.
AC P29598;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAUF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer 344;
RX MEDLINE=92233409; PubMed=1568219;
RA Henderson B.R.; Tansey W.P.; Phillips S.M.; Ramshaw I.A.;
RA Keiford R.F.;
RT "Transcriptional and posttranscriptional activation of urokinase
RT plasminogen activator gene expression in metastatic tumor cells.";
RL Cancer Res. 52:2489-2496(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Rabbani S.A.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH
CC CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE
CC LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; X63434; CAA45028.1; -.
CC DR EMBL; X65651; CAA46601.1; -.
CC DR PIR; S24604; S18932.
CC DR HSSP; P00749; 1KDU.
CC DR MEROPS; S01.231; -.
CC DR InterPro; IPR001314; Chymotrypsin.
CC DR InterPro; IPR006209; EGF like.
CC DR InterPro; IPR006210; IEGF.
CC DR InterPro; IPR000001; Kringle.
CC DR InterPro; IPR001254; Ser_protease_Try.
CC DR Pfam; PF00051; kringle; 1.
CC DR Pfam; PF00089; trypsin; 1.
CC DR PRINTS; PR00722; CHYMOTRYPSIN.
CC DR PRINTS; PR00018; KRINGLE.
CC DR ProDom; PD000395; Kringle; 1.
CC DR SMART; SM00130; KR; 1.
CC DR SMART; SM00020; Tryp_Spc; 1.
CC DR PROSITE; PS00022; EGF_1; 1.
CC DR PROSITE; PS01186; EGF_2; FALSE NEG.
CC DR PROSITE; PS00021; KRINGLE_1; 1.
CC DR PROSITE; PS50070; KRINGLE_2; 1.
CC DR PROSITE; PS50240; TRYPSIN_DOM; 1.
CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
CC DR PROSITE; PS00135; TRYPSIN_SER; 1.

```

```

KW Plasminogen activation: Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 20 432 CHAIN A (BY SIMILARITY).
FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 179 432 CHAIN B (BY SIMILARITY).
FT DOMAIN 27 63 EGF-LIKE.
FT DOMAIN 70 151 KRINGLE.
FT DOMAIN 152 178 CONNECTING PEPTIDE.
FT DOMAIN 179 432 SERINE PROTEASE.
FT DISULFID 31 39 BY SIMILARITY.
FT DISULFID 33 51 BY SIMILARITY.
FT DISULFID 53 62 BY SIMILARITY.
FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
FT DISULFID 210 226 BY SIMILARITY.
FT DISULFID 218 289 BY SIMILARITY.
FT DISULFID 314 383 BY SIMILARITY.
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 373 401 BY SIMILARITY.
FT ACT_SITE 225 225 CHARGE RELAY SYSTEM.
FT ACT_SITE 276 276 CHARGE RELAY SYSTEM.
FT ACT_SITE 377 377 CHARGE RELAY SYSTEM.
FT CONFLICT 16 16 N -> H (IN REF. 2).
FT CONFLICT 24 24 E -> G (IN REF. 2).
FT CONFLICT 332 332 D -> N (IN REF. 2).
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match 79.6%; Score 406; DB 1; Length 432;
Best Local Similarity 79.5%; Pred. No. 1.7e-39;
Matches 70; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 1 KTCVGNHGFYRGKASTDTMGRPCLPWNSATVLOOTYHAHRSDALQGLGKHNYCRNPDN 60
DB 68 KTCVGNHGFYRGKASTDTMGRPCLPWNSATVLOOTYHAHRSDALQGLGKHNYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 128 QRRPWCYVQIGLKQFVQECMVQDCSLK 155

RESULT 5
UROK_BOVIN STANDARD; PRT; 433 AA.
AC Q05589; Q28209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLA0.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic endothelium;
RX MEDLINE=93216119; PubMed=8385052;
RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,
RA Schieuning W.-D.;
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid.";
RL Gene 125:177-183(1993).
RP [2]
RW SEQUENCE OF 12-433 FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA.";
RL Int. Dairy J. 5:605-617(1995).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in

```

```

CC plasminogen to form plasmin.
CC -!- INDUCTION: By retinoic acid.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L03546; AAA51419.1; -.
DR EMBL; X85801; CAA59796.1; -.
DR PIR; JN0560; JN0560.
DR HSP; P00749; 1LMW.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle_1.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; Tryp_Spc; 1.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 180 CONNECTING PEPTIDE.
FT DOMAIN 181 433 SERINE PROTEASE.
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 189 189 A -> T (IN REF. 2).
SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA7027A CRC64;

Query Match 77.6%; Score 396; DB 1; Length 433;
Best Local Similarity 76.1%; Pred. No. 2.4e-36;
Matches 67; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

```

```

QY 1 KTCVGNHGFYRGKASTDTMGRPCLPWNSATVLOOTYHAHRSDALQGLGKHNYCRNPDN 60
DB 70 KTCVGNHGFYRGKASTDTMGRPCLPWNSATVLOOTYHAHRSDALQGLGKHNYCRNPDN 129

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 130 QRRPWCYVQIGLKQFVQECMVQDCSVK 157

```



```
RESULT 6
ID UROK MOUSE STANDARD; PRT; 433 AA.
AC P06869;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
DE PLAU
GN Mus musculus (Mouse).
OS
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. PubMed=2985383;
RA Belin D., Vassalli J.-D., Conbepine C., Godeau F., Nagamine Y.,
RA Reich E., Kocher H.P., Duvoisin R.M.;
RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
RT mouse urokinase-type plasminogen activator.";
RL Eur. J. Biochem. 148:225-232 (1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88163489; PubMed=2831940;
RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
RT "The murine urokinase-type plasminogen activator gene.";
RL Biochemistry 26:8270-8279 (1987).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR ENBL; X02389; CAA26231.1; -.
DR EMBL; M17922; AAA40539.1; -.
DR PIR; A29420; UKMS.
DR HSP; P00749; IKDU.
DR MEROPS; S01.231; -.
DR MGD; MGI:97611; Plau.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle 1.
DR Pfam; PF00089; trypsin 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE NEG.
```

```
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 178 CHAIN A (BY SIMILARITY).
FT CHAIN 157 178 SHORT A CHAIN (A1)
FT CHAIN 180 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 28 64 EGF-LIKE.
FT DOMAIN 71 152 KRINGLE.
FT DOMAIN 153 179 CONNECTING PEPTIDE.
FT DOMAIN 180 433 SERINE PROTEASE.
FT DISULFID 32 40 BY SIMILARITY.
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 54 63 BY SIMILARITY.
FT DISULFID 169 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM.
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
SQ SEQUENCE 433 AA; 48268 MW; A99C35F6250443F9 CRC64;

Query Match 75.9%; Score 387; DB 1; Length 433;
Best Local Similarity 72.7%; Pred. No. 2.6e-37;
Matches 64; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

Qy 1 KTCVGNHGPHYRGKASTDTMGRECLPWNSTVLQOTVHAHRSALQGLGKHNYCRNPDN 60
Db 69 KTCYHGNGOSYRGKANTDTKGRCLAWNAPVLQKPYNAHRPDAISLGKHKNYCRNPDN 128
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 129 QKRPWCYVQIGLRQFVQECMVHDCSLK 156

RESULT 7
URTB DESRO STANDARD; PRT; 431 AA.
AC P98121;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
DE beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237 (1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403 (1992).
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
```





RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toohyuk S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [9]  
RP SEQUENCE OF 212-361 FROM N.A.  
RX MEDLINE=83169656; PubMed=6572897;  
RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,  
RA Josephson S.;  
RT "Isolation of cDNA sequences coding for a part of human tissue  
RT plasminogen activator.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).  
RN [10]  
RP SEQUENCE OF 1-36 FROM N.A.  
RX MEDLINE=85289338; PubMed=3161893;  
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,  
RA Schleuning W.-D.;  
RT "Isolation and characterization of the human tissue-type plasminogen  
RT activator structural gene including its 5' flanking region.";  
RL J. Biol. Chem. 260:11223-11230(1985).  
RN [11]  
RP SEQUENCE OF 31-562 FROM N.A.  
RX MEDLINE=91291340; PubMed=1368681;  
RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;  
RT "Purification and characterization of tissue plasminogen activator  
RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";  
RL Agric. Biol. Chem. 55:1225-1232(1991).  
RN [12]  
RP SEQUENCE OF 36-562.  
RC TISSUE=Melanoma;  
RX MEDLINE=85000468; PubMed=6433976;  
RA Pohl G., Kaellstroem M., Bergsdorf N., Wallen P., Joernvall H.;  
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly  
RT derived amino acid sequence, identify the active site serine residue,  
RT establish glycosylation sites, and localize variant differences.";  
RL Biochemistry 23:3701-3707(1984).  
RN [13]  
RP SEQUENCE OF 33-52 AND 311-330.  
RC TISSUE=Melanoma;  
RX MEDLINE=83209620; PubMed=6682760;  
RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;  
RT "Purification and characterization of a melanoma cell plasminogen  
RT activator.";  
RL Eur. J. Biochem. 132:681-686(1983).  
RN [14]  
RP STRUCTURE OF CARBOHYDRATES.  
RX MEDLINE=90092112; PubMed=25131186;  
RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;  
RT "Carbohydrate structure of recombinant human uterine tissue  
RT plasminogen activator expressed in mouse epithelial cells.";  
RL Eur. J. Biochem. 186:273-286(1989).  
RN [15]  
RP CARBOHYDRATE-LINKAGE SITE THR-96.  
RX MEDLINE=91159408; PubMed=1900431;  
RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;

"Tissue plasminogen activator has an O-linked fucose attached to  
threonine-61 in the epidermal growth factor domain.";  
Biochemistry 30:2311-2314(1991).  
[16]  
RP DISULFIDE BONDS IN KRINGLE 2.  
RX MEDLINE=91244765; PubMed=1645336;  
RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;  
RT "Disulfide pairing of the recombinant kringle-2 domain of tissue  
RT plasminogen activator produced in *Escherichia coli*.";  
RL J. Biol. Chem. 266:10070-10072(1991).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.  
RX MEDLINE=96200385; PubMed=8613982;  
RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,  
RA Bode W.;  
RT "The 2.3 A crystal structure of the catalytic domain of recombinant  
RT two-chain human tissue-type plasminogen activator.";  
RL J. Mol. Biol. 258:117-135(1996).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.  
RX MEDLINE=97449126; PubMed=9305622;  
RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,  
RA Bode W.;  
RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray  
RT crystal structure of single-chain human tPA.";  
RL EMBO J. 16:4797-4805(1997).  
RN [19]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.  
RX MEDLINE=92118803; PubMed=1310033;  
RA de Vos A., Ullsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,  
RA Westbrook M.L., Kossiakof A.A.;  
RT "Crystal structure of the kringle 2 domain of tissue plasminogen  
RT activator at 2.4-A resolution.";  
RL Biochemistry 31:270-279(1992).  
RN [20]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RX MEDLINE=90122799; PubMed=2558718;  
RA Byeon I.-J.L., Kelley R.F., Llinas M.;  
RT "1H NMR structural characterization of a recombinant kringle 2 domain  
RT from human tissue-type plasminogen activator.";  
RL Biochemistry 28:9350-9360(1989).  
RN [21]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RX MEDLINE=91200042; PubMed=1901789;  
RA Byeon I.-J.L., Kelley R.F., Llinas M.;  
RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR  
RT assignments and secondary structure.";  
RL Eur. J. Biochem. 197:155-165(1991).  
RN [22]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RX MEDLINE=92106329; PubMed=1762144;  
RA Byeon I.-J.L., Llinas M.;  
RT "Solution structure of the tissue-type plasminogen activator kringle  
RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic  
RT drug.";  
RL J. Mol. Biol. 222:1035-1051(1991).  
RN [23]

Query Match 44.3%; Score 226; DB 1; Length 562;  
Best Local Similarity 47.7%; Pred. No. 1.4e-18;  
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCVEGNHGYRGKASTDTMGRECLPWN SATVLQOQTYHAHRSALQIGLGHNYCRNPDR 61  
Db 126 TCYEDQGISYRGTSWTASGAECTWNSSALAKQPYSGRRPDAIRLGLGNHNYCRNPDR 185

QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87  
Db 186 SRPWCYVFKAGKYSSEFCSTPACSEG 211

RESULT 10  
TPA\_RAT

ID	TPA_RAT	STANDARD;	PRT;	559 AA.
AC	P19637;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)			
DE	(t-PA) (t-plasminogen activator).			
GN	PLAT.			
OS	Rattus norvegicus (Rat).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OC	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89170114; PubMed=3148445;			
RA	NY T., Leonardson G., Hsueh A.J.W.;			
RA	"Cloning and characterization of a cDNA for rat tissue-type			
RT	plasminogen activator.";			
RL	DNA 7:671-677(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90130448; PubMed=2105315;			
RA	Feng P., Ohlsson M., Ny T.;			
RA	"The structure of the TATA-less rat tissue-type plasminogen activator			
RT	gene. Species-specific sequence divergences in the promoter predict			
RT	differences in regulation of gene expression.";			
RL	J. Biol. Chem. 265:2022-2027(1990).			
CC	-1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMAGEN PLASMINOGEN			
CC	TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY			
CC	CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT			
CC	ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND			
CC	MANY OTHER PHYSIOPATHOLOGICAL EVENTS.			
CC	-1- CATALYTIC ACTIVITY: Specific cleavage of Arg -Val bond in			
CC	plasminogen to form plasmin.			
CC	-1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE			
CC	BOND.			
CC	-1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.			
CC	-1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER			
CC	PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER			
CC	ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.			
CC	-1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A			
CC	CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-1- SIMILARITY: Contains 1 EGF-like domain.			
CC	-1- SIMILARITY: Contains 1 fibronectin type I domain.			
CC	-1- SIMILARITY: Contains 2 kringle domains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collabora-			
CC	tion between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no			
CC	modified and this statement is not removed. Usage by and for commer-			
CC	cial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announ">http://www.isb-sib.ch/announ</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; M23697; AAA41812.1; -			
DR	EMBL; M31197; AAA42261.1; -			
DR	EMBL; M31185; AAA42261.1; JOINED.			
DR	EMBL; M31186; AAA42261.1; JOINED.			
DR	EMBL; M31187; AAA42261.1; JOINED.			
DR	EMBL; M31188; AAA42261.1; JOINED.			
DR	EMBL; M31189; AAA42261.1; JOINED.			
DR	EMBL; M31190; AAA42261.1; JOINED.			
DR	EMBL; M31191; AAA42261.1; JOINED.			
DR	EMBL; M31192; AAA42261.1; JOINED.			
DR	EMBL; M31193; AAA42261.1; JOINED.			
DR	EMBL; M31194; AAA42261.1; JOINED.			
DR	EMBL; M31195; AAA42261.1; JOINED.			
DR	EMBL; M31196; AAA42261.1; JOINED.			
DR	EMBL; A19618; CAA01482.1; -			
DR	PIR; A35029; A35029.			
DR	HSSP; P00750; LRTE.			
DR	MEROPS; S01_232; -			



RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN  
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY  
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT  
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND  
 CC MANY OTHER PHYSIOLOGICAL EVENTS.  
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE  
 CC BOND.  
 CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.  
 CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER  
 CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER  
 CC ARG-308 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FACTOR XA.  
 CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A  
 CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -1- SIMILARITY: Contains 2 kringle domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J03520; AAA40470.1; --  
 DR EMBL; BC011256; AAH11256.1; --  
 DR PIR; A29941; A29941.  
 DR HSP; P00750; 1A5H.  
 DR MEROPS; S01-232; --  
 DR MGD; MGI:97610; Plat.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR006209; EGF like.  
 DR InterPro; IPR000083; Fibrinctn1.  
 DR InterPro; IPR006210; tEGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Ser-protease\_Try.  
 DR Pfam; PF00008; EGF; 1.  
 DR Pfam; PF00039; fn1; 1.  
 DR Pfam; PF00051; kringle; 2.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 2.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00058; FN1; 1.  
 DR SMART; SM00130; KR; 2.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 2.  
 DR PROSITE; PS50070; KRINGLE\_2; 2.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 FT SIGNAL 1 17  
 FT PROPEP 18 29  
 FT CHAIN 30 559  
 FT CHAIN 30 308  
 FT CHAIN 309 559  
 FT CHAIN 309 559

FT DOMAIN 36 78  
 FT DOMAIN 79 117  
 FT DOMAIN 124 205  
 FT DOMAIN 213 294  
 FT DOMAIN 213 294  
 FT ACT\_SITE 359 355  
 FT ACT\_SITE 355 355  
 FT ACT\_SITE 404 404  
 FT ACT\_SITE 510 510  
 FT DISULFID 38 68  
 FT DISULFID 66 75  
 FT DISULFID 83 94  
 FT DISULFID 88 105  
 FT DISULFID 107 116  
 FT DISULFID 124 205  
 FT DISULFID 145 187  
 FT DISULFID 176 200  
 FT DISULFID 213 294  
 FT DISULFID 234 276  
 FT DISULFID 265 289  
 FT DISULFID 297 428  
 FT DISULFID 340 356  
 FT DISULFID 348 417  
 FT DISULFID 442 516  
 FT DISULFID 474 490  
 FT DISULFID 506 534  
 FT CARBOHYD 149 149  
 FT CARBOHYD 481 481  
 FT CONFLICT 260 260  
 FT CONFLICT 325 325  
 SQ SEQUENCE 559 AA; 63122 MW; 8CCBE2BDB94514D9 CRC64;  
 Query Match 42.9%; Score 219; DB 1; Length 559;  
 Best Local Similarity 46.0%; Pred. No. 9.3e-18;  
 Matches 40; Conservative 11; Mismatches 36; Indels 0; Gaps 0;  
 QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHSDALQLGLGHKHYCRNPDNR 61  
 DB 123 TCPEEQGITRYGTWSTAESGAECINWNSSVLSPKPNARPNALKLGHNHYCRNPDNR 182  
 QY 62 RRPWCYVQVGLKPLVQECMVHDCADGK 88  
 DB 183 LKPCWYVFKAGKYTTEFCSTPACPKGK 209  
 RESULT 13  
 ID URT1 DESRO STANDARD; PRT; 477 AA.  
 AC P98119;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA  
 DE alpha-1).  
 OS Desmodus rotundus (Vampire bat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;  
 OC Desmodontinae; Desmodus.  
 OX NCBI\_TaxID=9430;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Salivary gland;  
 RX MEDLINE=92039036; PubMed=1937019;  
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
 RA Alagon A., Donner P., Schleuning W.D.;  
 RT "The plasminogen activator family from the salivary gland of the  
 RT vampire bat Desmodus rotundus: cloning and expression.";  
 RL Gene 105:229-237(1991).  
 RN [2]  
 RP CHARACTERIZATION.  
 RX MEDLINE=93393059; PubMed=1309059;  
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,  
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,

```

RA  Donner P.;
RT  "Plasminogen activators from the saliva of Desmodus rotundus (common
RL  vampire bat): unique fibrin specificity.";
RN  Ann. N.Y. Acad. Sci. 667:395-403(1992).
RC  [3]
RD  X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RE  TISSUE=salivary gland;
RF  MEDLINE=9802741; PubMed=9354616;
RG  Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
RH  Schleuning W.D., Bode W.;
RI  "Catalytic domain structure of vampire bat plasminogen activator: a
RJ  molecular paradigm for proteolysis without activation cleavage.";
RK  Biochemistry 36:13483-13493(1997).
RL  !- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
RM  EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. POTENT THROMBOLYTIC AGENT.
RN  !- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
RP  plasminogen to form plasmin.
RQ  !- ENZYME REGULATION: ACTIVATION OF ARG-|-VAL BOND IN
RS  THE PRESENCE OF FIBRIN I.
RT  !- SUBUNIT: Monomer.
RU  !- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
RV  AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
RW  STIMULATION OF ACTIVITY.
RX  !- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
RY  !- SIMILARITY: Contains 1 EGF-like domain.
RZ  !- SIMILARITY: Contains 1 fibronectin type I domain.
RA  !- SIMILARITY: Contains 1 kringle domain.
RB  -----
RC  This SWISS-PROT entry is copyright. It is produced through a collaboration
RD  between the Swiss Institute of Bioinformatics and the EMBL outstation -
RE  the European Bioinformatics Institute. There are no restrictions on its
RF  use by non-profit institutions as long as its content is in no way
RG  modified and this statement is not removed. Usage by and for commercial
RH  entities requires a license agreement (See http://www.isb-sib.ch/announce/
RI  or send an email to license@isb-sib.ch).
RJ  -----
RK  EMBL; M63987; AAA31591.1; -.
RL  EMBL; M63986; AAA31592.1; -.
RM  PIR; JS0597; JS0597.
RN  PDB; 1A51; 23-MAR-99.
RO  MEROPS; S01.232; -.
RP  Glycosultide; P98119; -.
RQ  InterPro; IPR001314; Chymotrypsin.
RR  InterPro; IPR006209; EGF like.
RS  InterPro; IPR000083; Fibrinctnl.
RT  InterPro; IPR006210; IEGF.
RU  InterPro; IPR000001; Kringle.
RV  InterPro; IPR001254; Ser_protease_Try.
RW  Pfam; PF00008; EGF; 1.
RX  Pfam; PF00039; fnl; 1.
RY  Pfam; PF00051; kringle; 1.
RZ  Pfam; PF00089; trypsin; 1.
RA  PRINTS; PR00722; CHYMOTRYPSIN.
RB  PRINTS; PR00018; KRINGLE.
RC  ProDom; PD000395; Kringle; 1.
RD  SMART; SM00181; EGF; 1.
RE  SMART; SM00058; FNI; 1.
RF  SMART; SM00130; KR; 1.
RG  SMART; SM00020; Tryp_SPC; 1.
RH  PROSITE; PS00022; EGF_1; 1.
RI  PROSITE; PS01186; EGF_2; 1.
RJ  PROSITE; PS01253; FIBRONECTIN_1; 1.
RK  PROSITE; PS00021; KRINGLE_1; 1.
RL  PROSITE; PS50070; KRINGLE_2; 1.
RM  PROSITE; PS50240; TRYPSIN_DOM; 1.
RN  PROSITE; PS00134; TRYPSIN_HIS; 1.
RO  PROSITE; PS00135; TRYPSIN_SER; 1.
RP  Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
RQ  Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.
RR  SIGNAL 1 36 POTENTIAL.
RS  CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 1.
RT  DOMAIN 40 82 FIBRONECTIN TYPE-I.
RU  DOMAIN 83 121 EGF-LIKE.

```

```

FT  DOMAIN 128 209 KRINGLE.
FT  DOMAIN 225 477 SERINE PROTEASE.
FT  ACT_SITE 272 272 CHARGE RELAY SYSTEM.
FT  ACT_SITE 321 321 CHARGE RELAY SYSTEM.
FT  ACT_SITE 428 428 CHARGE RELAY SYSTEM.
FT  DISULFID 42 72 BY SIMILARITY.
FT  DISULFID 70 79 BY SIMILARITY.
FT  DISULFID 87 98 BY SIMILARITY.
FT  DISULFID 92 109 BY SIMILARITY.
FT  DISULFID 111 120 BY SIMILARITY.
FT  DISULFID 128 209 BY SIMILARITY.
FT  DISULFID 149 191 BY SIMILARITY.
FT  DISULFID 180 204 BY SIMILARITY.
FT  DISULFID 214 345 BY SIMILARITY.
FT  DISULFID 257 273 BY SIMILARITY.
FT  DISULFID 265 334 BY SIMILARITY.
FT  DISULFID 359 434 BY SIMILARITY.
FT  DISULFID 391 407 BY SIMILARITY.
FT  DISULFID 424 452 BY SIMILARITY.
FT  CARBOHYD 153 153 N-LINKED (GLCNAC. .).
FT  CARBOHYD 398 398 /FTid=CAR_000027.
FT  CARBOHYD 398 398 N-LINKED (GLCNAC. .).
FT  CARBOHYD 398 398 /FTid=CAR_000028.
FT  TURN 214 215
FT  STRAND 223 224
FT  TURN 226 227
FT  STRAND 230 231
FT  HELIX 234 236
FT  TURN 238 239
FT  STRAND 240 245
FT  STRAND 254 263
FT  TURN 264 265
FT  STRAND 266 269
FT  HELIX 271 273
FT  TURN 280 282
FT  STRAND 284 287
FT  TURN 297 298
FT  STRAND 300 309
FT  TURN 311 312
FT  TURN 315 317
FT  TURN 319 320
FT  STRAND 323 328
FT  STRAND 338 338
FT  TURN 339 340
FT  STRAND 341 341
FT  STRAND 345 345
FT  TURN 349 350
FT  TURN 355 356
FT  STRAND 358 363
FT  STRAND 366 366
FT  STRAND 374 374
FT  STRAND 379 385
FT  HELIX 388 390
FT  TURN 393 398
FT  TURN 403 404
FT  STRAND 405 409
FT  TURN 425 426
FT  TURN 428 429
FT  STRAND 431 436
FT  TURN 437 438
FT  STRAND 439 448
FT  TURN 455 456
FT  STRAND 459 463
FT  HELIX 464 467
FT  HELIX 468 474
SQ  SEQUENCE 477 AA; 53616 MW; AA06FD1739C10B5E CRC64;

```

```

Query Match 41.8%; Score 213; DB 1; Length 477;
Best Local Similarity 46.4%; Pred. No. 3.9e-17;
Matches 39; Conservative 10; Mismatches 35; Indels 0; Gaps 0;
2 TCYEGNGHFYRGKASTDTMGRCPLPWNSATVLQOQYHAHRSDALQLGLGKHNYCRNPDR 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
: |||: : |||: : |||: : |||: : |||: : |||: : |||: : |||: : |||: :

```



```
Db 127 TCYEGQGVTVGTWSTAESRVECEINWNSSLLTRTYNGRMPDAPNLGLGHNNHYCRNPNGA 186
QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
Db 187 PKWCYVVKAGKFTSESCSVPCS 210

RESULT 14
TPA_BOVIN
ID TPA_BOVIN STANDARD; PRT; 566 AA.
AC Q28198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RL Int. Dairy J. 5:605-617(1995).."
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 2 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X85800; CAA59795.1; -.
CC HSP; P00750; IRTF.
CC MERO; S01.232; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR000083; Fibinctnl.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00039; fnl; 1.
CC Pfam; PF00051; kringle; 2.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 2.
```

```
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FNI; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01185; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50070; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT PROPEP 22 33 BY SIMILARITY.
FT CHAIN 34 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 34 314 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 315 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT CHAIN 315 566 CHAIN.
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.
FT DOMAIN 83 121 EGF-LIKE.
FT DOMAIN 128 209 KRINGLE 1.
FT DOMAIN 219 300 KRINGLE 2.
FT DOMAIN 315 566 SERINE PROTEASE.
FT ACT_SITE 361 361 CHARGE RELAY SYSTEM.
FT ACT_SITE 410 410 CHARGE RELAY SYSTEM.
FT ACT_SITE 517 517 CHARGE RELAY SYSTEM.
FT DISULFID 42 72 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
FT DISULFID 149 191 BY SIMILARITY.
FT DISULFID 180 204 BY SIMILARITY.
FT DISULFID 219 300 BY SIMILARITY.
FT DISULFID 240 282 BY SIMILARITY.
FT DISULFID 271 295 BY SIMILARITY.
FT DISULFID 303 434 INTERCHAIN (BY SIMILARITY).
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 354 423 BY SIMILARITY.
FT DISULFID 448 523 BY SIMILARITY.
FT DISULFID 480 496 BY SIMILARITY.
FT DISULFID 513 541 BY SIMILARITY.
FT CARBOHYD 153 153 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BEB4E32276C3 CRC64;

Query Match 41.0%; Score 209; DB 1; Length 566;
Best Local Similarity 47.6%; Pred. No. 1.3e-16;
Matches 39; Conservative 9; Mismatches 34; Indels 0; Gaps 0;

Oy 3 CYEGNGHFYRGKASTDTMGPRCLPWN SATVLPQYTHAHRSDALQGLGKHNYCRNPDRR 62
Db 219 CYTNGLAYRGTSHYKSGASCLPWN SVFLTSKIYTAWKSAPALGLGKHNYCRNPQDA 278

Oy 63 RPPCYVQVGLKPLVQECMVHDC 84
Db 279 QPWCHVWKDRQLTWECVDPQC 300

RESULT 15
UROK_CHICK
ID UROK_CHICK STANDARD; PRT; 434 AA.
AC P15120;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
DE Gallus gallus (Chicken).
OS Gallus gallus (Chicken).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RX MEDLINE=90110185; PubMed=2295632;  
RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;  
RT "The chicken urokinase-type plasminogen activator gene.";  
RL J. Biol. Chem. 265:1339-1344(1990).  
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: Contains 1 kringle domain.  
CC -1- SIMILARITY: Contains 1 EGF-like domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J05187; AAA49131.1; -;  
DR EMBL; J05188; AAA49130.1; -;  
DR PIR; A35005; A35005.  
DR HGSP; P00763; IDPO.  
DR MEROPS; S01.231; -;  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Ser.protease\_Try.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Signal; Zymogen.  
FT SIGNAL 1 20  
FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
FT CHAIN 21 171 CHAIN A (BY SIMILARITY).  
FT CHAIN 173 434 CHAIN B (BY SIMILARITY).  
FT DOMAIN 36 72 EGF-LIKE.  
FT DOMAIN 79 158 KRINGLE.  
FT DOMAIN 159 172 CONNECTING PEPTIDE.  
FT DOMAIN 173 434 SERINE PROTEASE.  
FT DISULFID 40 48 BY SIMILARITY.  
FT DISULFID 42 60 BY SIMILARITY.  
FT DISULFID 62 71 BY SIMILARITY.  
FT DISULFID 162 296 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 202 218 BY SIMILARITY.  
FT DISULFID 210 285 BY SIMILARITY.  
FT DISULFID 310 379 BY SIMILARITY.  
FT DISULFID 342 358 BY SIMILARITY.  
FT DISULFID 369 397 BY SIMILARITY.  
FT ACT\_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 373 373 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT CARBOHYD 228 228 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 434 AA; 49400 MW; BD861048DD666A55 CRC64;

Query Match 38.9%; Score 198.5; DB 1; Length 434;  
Best Local Similarity 54.4%; Pred. No. 1.7e-15;  
Matches 37; Conservative 7; Mismatches 19; Indels 5; Gaps 2;  
QY 3 CYEGNGHFYRGKASTDTMGPCLPWNSATVLQ-QTYHAHRSDALQLGLGKHNYCRNPDR 61  
DB 79 CYSNGGDIYRGMAEDP---GCLYWDHPSVIRWGDYHADLKNALQLGLGKHNYCRNPGR 134  
QY 62 RRPWCYVQ 69  
DB 135 SRPWCYTK 142

## RESULT 16

HGFA\_HUMAN STANDARD; PRT; 655 AA.  
AC Q04756; Q14726;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF  
DE activator) (HGFA).  
GN HGFA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Liver, and Serum;  
RX MEDLINE=93252878; PubMed=7683665;  
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,  
RA Kitamura N.;  
RT "Molecular cloning and sequence analysis of the cDNA for a human  
RT serine protease responsible for activation of hepatocyte growth  
RT factor. Structural similarity of the protease precursor to blood  
RT coagulation factor XII.";  
RL J. Biol. Chem. 268:10024-10028(1993).  
RN [2]  
RP SEQUENCE OF 40-655 FROM N.A.  
RA Zhao S., Odell C.;  
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY  
CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.  
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A  
CC DISULFIDE BOND.  
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN  
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.  
CC -1- TISSUE SPECIFICITY: LIVER.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: Contains 2 EGF-like domains.  
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -1- SIMILARITY: Contains 1 fibronectin type II domain.  
CC -1- SIMILARITY: Contains 1 kringle domain.  
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D14012; BAA03113.1; -;  
DR EMBL; Z69923; CAA93803.1; -;  
DR PIR; A46688; A46688.  
DR HGSP; P00763; IDPO.  
DR MEROPS; S01.228; -;  
DR Genew; HGNC:4894; HGFA.  
DR MIM; 604552; -;  
DR GO; GO:0005576; C:extracellular; TAS.



```
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringl.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringl; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringl; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringl; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29 BY SIMILARITY.
FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT CHAIN.
FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG CHAIN.
FT DOMAIN 105 145 FIBRONECTIN TYPE-II.
FT DOMAIN 157 195 EGF-LIKE 1.
FT DOMAIN 197 237 FIBRONECTIN TYPE-I.
FT DOMAIN 238 276 EGF-LIKE 2.
FT DOMAIN 283 364 KRINGLE.
FT DOMAIN 406 653 SERINE PROTEASE.
FT ACT_SITE 445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 105 130 BY SIMILARITY.
FT DISULFID 119 145 BY SIMILARITY.
FT DISULFID 161 172 BY SIMILARITY.
FT DISULFID 166 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 199 227 BY SIMILARITY.
FT DISULFID 225 234 BY SIMILARITY.
FT DISULFID 242 253 BY SIMILARITY.
FT DISULFID 247 264 BY SIMILARITY.
FT DISULFID 266 275 BY SIMILARITY.
FT DISULFID 283 364 BY SIMILARITY.
FT DISULFID 304 346 BY SIMILARITY.
FT DISULFID 335 359 BY SIMILARITY.
FT DISULFID 392 519 INTERCHAIN (BY SIMILARITY).
FT DISULFID 430 446 BY SIMILARITY.
FT DISULFID 438 508 BY SIMILARITY.
FT DISULFID 533 602 BY SIMILARITY.
FT DISULFID 565 581 BY SIMILARITY.
FT DISULFID 592 620 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 164 164 G -> W (IN REF. 2).
SQ SEQUENCE 653 AA; 70567 MW; 38B4B20255DF7FDC CRC64;
```

```
Query Match 36.9%; Score 188; DB 1; Length 653;
Best Local Similarity 53.0%; Pred No. 4.2e-14;
Matches 35; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

Qy 3 CYEGNGHFYRGKASTDTMGPCLPWNSATVLTQYTHAHRSDALQLGLGKHNCRPNRR 62
   : ||| ||| ||| ||| ||| : : : ||| ||| |||
Db 283 CFLGNGTEYRGVASTAASGLSLAUNSDLLVQLHVDVAAVLLGLGPHAYCRNPKDE 342
   : ||| ||| ||| ||| ||| : : : ||| ||| |||

Qy 63 RWCYV 68
   |||||
Db 343 RWCYV 348

RESULT 18
FA12_CAVPO
ID FA12_CAVPO STANDARD; PRT; 603 AA.
AC Q04962;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF) (Fragment).
GN F12.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
RC TISSUE=Liver;
RX MEDLINE=93003367; PubMed=1390917;
RA Samba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
RA Kambara T., Okabe H.;
RT "Primary structure of guinea-pig Hageman factor: sequence around the
RT cleavage site differs from the human molecule.";
RL Biochim. Biophys. Acta 1159:113-121(1992).
CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-Ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X68615; CAA48600.1; -.
CC PIR; S28941; S28941.
CC HSP; P00763; IDPO.
CC MEROPS; S01.211; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR000083; Fibrinctnl.
CC InterPro; IPR000562; FN_type_II.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringl.
CC InterPro; IPR001254; Ser_protease_Try.
```

```

DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
FT NON_TER 1 1
FT SIGNAL <1 38
FT CHAIN 19 158
FT CHAIN 359 603
FT DOMAIN 46 87
FT DOMAIN 93 130
FT DOMAIN 132 172
FT DOMAIN 173 209
FT DOMAIN 216 294
FT DOMAIN 312 342
FT DOMAIN 359 603
FT ACT_SITE 398 398
FT ACT_SITE 447 447
FT ACT_SITE 551 551
FT DISULFID 97 109
FT DISULFID 103 118
FT DISULFID 120 129
FT DISULFID 134 162
FT DISULFID 160 169
FT DISULFID 177 188
FT DISULFID 182 197
FT DISULFID 199 208
FT DISULFID 216 294
FT DISULFID 237 276
FT DISULFID 265 289
FT DISULFID 345 472
FT DISULFID 383 399
FT DISULFID 391 461
FT DISULFID 422 425
FT DISULFID 488 557
FT DISULFID 520 536
FT DISULFID 547 578
FT CARBOHYD 248 248
FT CARBOHYD 270 270
FT CARBOHYD 419 419
SQ SEQUENCE 603 AA; 66795 MW; 48DC6B946FB9ED59 CRC64;

Query Match 35.4%; Score 180.5; DB 1; Length 603;
Best Local Similarity 42.9%; Pred. No. 2.8e-13;
Matches 36; Conservative 11; Mismatches 32; Indels 5; Gaps 2;
QY 2 TCYEGHGFYRGKASTDTMGSPCLPFWNSATVLQOYTHAHRSD-ALQLGLGKHNYCRNPDN 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
215 SCYEGRGVSYRGWARTTYSGAKCQENAS----EATYRNWTAEQALRGHHTFCRNP 270
Db
QY 61 RRRPWCYVQVGLKPLVQPCWVHDC 84
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 271 DTRPWCFTWMGNRLSWEYCDLAQC 294
RESULT 19
FA12_HUMAN
ID FA12_HUMAN STANDARD; PRT; 615 AA.
AC P00748; P78339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF).
GN F12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. PubMed=2888762;
RX MEDLINE=88007593; PubMed=2888762;
RA Cool D.E., McGillivray R.T.A.;
RT "Characterization of the human blood coagulation factor XII gene.
RT Intron/exon gene organization and analysis of the 5'-flanking
RT region."
RL J. Biol. Chem. 262:13662-13673(1987).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
RA Rajkumar N., Toch E.J., Yi Q., Nickerson D.A.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 4-615 FROM N.A.
RX MEDLINE=86176794; PubMed=3754331;
RA Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA Cortese R.;
RT "cDNA sequence coding for human coagulation factor XII (Hageman)."
RL Nucleic Acids Res. 14:3146-3146(1986).
RN [4]
RP SEQUENCE OF 14-615 FROM N.A.
RX MEDLINE=86033830; PubMed=3877053;
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA McGillivray R.T.A.;
RT "Characterization of human blood coagulation factor XII cDNA.
RT Prediction of the primary structure of factor XII and the tertiary
RT structure of beta-factor XIIa."
RL J. Biol. Chem. 260:13666-13676(1985).
RN [5]
RP SEQUENCE OF 146-615 FROM N.A.
RX MEDLINE=86216049; PubMed=3011063;
RA Que B.G., Davie E.W.;
RT "Characterization of a cDNA coding for human factor XII (Hageman
RT factor)."
RL Biochemistry 25:1525-1528(1986).
RN [6]
RP SEQUENCE OF 20-379.
RX MEDLINE=85182674; PubMed=3886654;
RA McMullen B.A., Fujikawa K.;
RT "Amino acid sequence of the heavy chain of human alpha-factor XIIa
RT (activated Hageman factor)."
RL J. Biol. Chem. 260:5328-5341(1985).
RN [7]
RP SEQUENCE OF 354-362 AND 373-615.
RX MEDLINE=83291041; PubMed=6604055;
RA Fujikawa K., McMullen B.A.;
RT "Amino acid sequence of human beta-factor XIIa."
RL J. Biol. Chem. 258:10924-10933(1983).
RN [8]
RP SEQUENCE OF 561-615 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=96133302; PubMed=8528215;
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;
RT "The novel acceptor splice site mutation 11396(G-->A) in the factor
RT XII gene causes a truncated transcript in cross-reacting material

```



```
RESULT 20
APOA_MACMU
ID APOA_MACMU STANDARD; PRT; 1420 AA.
AC P1417;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).
GN Lp(a).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=99174660; PubMed=2925643;
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT synthesis.";
RL J. Biol. Chem. 264:5957-5965(1989).
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin (by similarity).
CC -!- PTM: N- and O-glycosylated (by similarity).
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringle IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrin(ogen) binding.
CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC leading to the formation of the so called mini-Lp(a). Apo(a)
CC fragments accumulate in atherosclerotic lesions, where they may
CC promote thrombogenesis. O-glycosylation may limit the extent of
CC proteolytic fragmentation (by similarity).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: Contains at least 10 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; J04635; AAA36833.1; -
CC PIR; A32869; A32869.
CC HSP; P00747; 2PK4.
CC
CC MEROPS; S01.226; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser. protease_Try.
CC Pfam; PF00051; kringle; 11.
CC Pfam; PF00089; trypsin.1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 10.
CC SMART; SM00130; KR; 10.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS00021; KRINGLE_1; 10.
CC PROSITE; PS50070; KRINGLE_2; 10.
CC PROSITE; PS0240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; FALSE NEG.
CC PROSITE; PS00135; TRYPSIN_SER; FALSE NEG.
CC Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
CC Kringle; Repeat; Atherosclerosis.
CC NON_TER 1
CC 49 127 KRINGLE 1.
CC 163 241 KRINGLE 2.
CC DOMAIN
FT
```

```
FT DOMAIN 277 355 KRINGLE 3.
FT DOMAIN 391 469 KRINGLE 4.
FT DOMAIN 505 583 KRINGLE 5.
FT DOMAIN 619 697 KRINGLE 6.
FT DOMAIN 725 803 KRINGLE 7.
FT DOMAIN 839 917 KRINGLE 8.
FT DOMAIN 953 1031 KRINGLE 9.
FT DOMAIN 1067 1145 KRINGLE 10.
FT DOMAIN 1191 1420 SERINE PROTEASE.
SQ SEQUENCE 1420 AA; 158367 MW; BE102949E03CB0E CRC64;

Query Match 31.6%; Score 161; DB 1; Length 1420;
Best Local Similarity 40.9%; Pred. No. 1.3e-10;
Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;

Oy 3 CYEGNGHFYRGKASTDTWGRPCLPWNSATVLQ--QTYHAHRS DALQLGLGKHNYCRNPDN 60
Db 1068 CYHGNGSYRGFTSTVTGTCQSSSWTPHQHKRTPEHPNDLTM-----NYCRNPD 1122

Oy 61 RRRPWCYVQVGLKPLVQE--CMVHDCAD 86
Db 1123 DTGFWCFT--MDPSVRREYCNLTRCSD 1147

RESULT 21
KRM1_MOUSE
ID KRM1_MOUSE STANDARD; PRT; 473 AA.
AC Q99N43;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.
RC TISSUE=Brain, and Kidney;
RX MEDLINE=21167372; PubMed=11267660;
RA Nakamura T., Aoki S., Kitajima K., Takahashi T., Matsumoto K.,
RA Nakamura T.;
RT "Molecular cloning and characterization of Kremen, a novel
RT kringle-containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: In the adult, widely expressed with high
CC levels in heart, lung, kidney, skeletal muscle and testis.
CC -!- DEVELOPMENTAL STAGE: In the embryo, expression is first detected
CC on day 9 and increases up to day 18. Lower levels are found in
CC ectodermal ridge (AER) of the developing fore- and hindlimb buds,
CC the telencephalon and the first brachial arch. At 10.5 dpc,
CC expression is also observed in the myotome and in sensory tissues
CC such as the nasal pit and optic vesicle.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AB059617; BAB40968.1; -
CC HSP; P00747; ICEA.
CC DOMAIN
DR
```

```

DR MGD; MGI:1933988; Kremen.
DR GO; GO:0016021; C: integral to membrane; NAS.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS001180; CUB; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00021; KRINGLE_2; 1.
DR PROSITE; PS00070; Kringle.
KW Wnt signaling pathway; Signal; Transmembrane; Kringle.
FT SIGNAL 1 19
FT CHAIN 20 473
FT DOMAIN 21 392
FT TRANSMEM 393 413
FT DOMAIN 414 473
FT DOMAIN 31 114
FT DOMAIN 120 210
FT DOMAIN 214 321
FT CARBOHYD 59 59
FT CARBOHYD 217 217
FT CARBOHYD 255 255
FT CARBOHYD 293 293
FT CARBOHYD 333 333
FT CARBOHYD 345 345
FT CARBOHYD 345 345
SQ SEQUENCE 473 AA; 5176 MW; 586827788BE3FDD1 CRC64;

Query Match 31.2%; Score 159; DB 1; Length 473;
Best Local Similarity 45.6%; Pred. No. 6.7e-11;
Matches 31; Conservative 7; Mismatches 26; Indels 4; Gaps 2;

QY 3 CYEGNGHYRGKASTDTM--GRPLPWSATVLQQTTHAHRSDALQLGLGKHNYCRNPDN 60
DB 32 CFTANGADYRGTSWTALQGGKPCLFWNE--TFQHPYNTLKYPNEGGLGHEHNYCRNPDG 89

QY 61 RRRPWCYV 68
DB 90 DVSPWCYV 97

RESULT 22
KRM1_RAT
ID KRM1_RAT STANDARD; PRT; 473 AA.
AC Q924S4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakamura T.; Nakamura T.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; AB065090; BAB62003.1; -.
CC GO; GO:0016021; C: integral to membrane; ISS.
CC InterPro; IPR000859; CUB_domain.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR002889; WSC.
CC Pfam; PF00431; CUB; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF01822; WSC; 1.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00042; CUB; 1.
CC SMART; SM00130; KR; 1.
CC PROSITE; PS001180; CUB; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00021; KRINGLE_2; 1.
CC PROSITE; PS00070; Kringle.
KW Wnt signaling pathway; Signal; Transmembrane; Kringle.
FT SIGNAL 1 19
FT CHAIN 20 473
FT DOMAIN 21 392
FT TRANSMEM 393 413
FT DOMAIN 414 473
FT DOMAIN 31 114
FT DOMAIN 120 210
FT DOMAIN 214 321
FT CARBOHYD 59 59
FT CARBOHYD 217 217
FT CARBOHYD 255 255
FT CARBOHYD 293 293
FT CARBOHYD 333 333
FT CARBOHYD 345 345
FT CARBOHYD 345 345
SQ SEQUENCE 473 AA; 51869 MW; 9B510857DF856F08 CRC64;

Query Match 31.2%; Score 159; DB 1; Length 473;
Best Local Similarity 45.6%; Pred. No. 6.7e-11;
Matches 31; Conservative 7; Mismatches 26; Indels 4; Gaps 2;

QY 3 CYEGNGHYRGKASTDTM--GRPLPWSATVLQQTTHAHRSDALQLGLGKHNYCRNPDN 60
DB 32 CFTANGADYRGTSWTALQGGKPCLFWNE--TFQHPYNTLKYPNEGGLGHEHNYCRNPDG 89

QY 61 RRRPWCYV 68
DB 90 DVSPWCYV 97

RESULT 23
KRM1_XENLA
ID KRM1_XENLA STANDARD; PRT; 452 AA.
AC Q90Y90;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakamura T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 CUB domain.

```



```

CC CC -!- SIMILARITY: Contains 1 kringle domain.
CC CC -!- SIMILARITY: Contains 1 WSC domain.
CC CC -----
CC CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AB070851; BAB64294.1; --
CC CC InterPro; IPR000859; CUB domain.
CC CC InterPro; IPR000001; Kringle.
CC CC InterPro; IPR002889; WSC.
CC CC Pfam; PF00431; CUB; 1.
CC CC Pfam; PF01822; WSC; 1.
CC CC PRINTS; PR00018; KRINGLE.
CC CC ProDom; PD000395; Kringle; 1.
CC CC SMART; SM00042; CUB; 1.
CC CC SMART; SM00130; KR; 1.
CC CC SMART; SM00321; WSC; 1.
CC CC PROSITE; PS01180; CUB; 1.
CC CC PROSITE; PS00021; KRINGLE_1; 1.
CC CC PROSITE; PS00070; KRINGLE_2; 1.
CC CC Wnt signaling pathway; Glycoprotein; Kringle;
CC KW Transmembrane, Signal.
CC FT SIGNAL 1 22
CC FT CHAIN 23 452
CC FT DOMAIN 23 369
CC FT TRANSMEM 370 390
CC FT DOMAIN 391 452
CC FT DOMAIN 29 112
CC FT DOMAIN 118 208
CC FT DOMAIN 212 319
CC FT CARBOHYD 43 43
CC FT CARBOHYD 57 57
CC FT CARBOHYD 215 215
CC FT CARBOHYD 253 253
CC FT CARBOHYD 291 291
CC FT CARBOHYD 328 328
CC FT CARBOHYD 344 344
CC SQ SEQUENCE 452 AA; 50188 MW; ED24BCD1AF4564E2 CRC64;
CC
CC Query Match 31.0%; Score 158; DB 1; Length 452;
CC Best Local Similarity 44.1%; Pred. No. 8.4e-11;
CC Matches 30; Conservative 9; Mismatches 25; Indels 4; Gaps 2;
CC
CC Qy 3 CYEGNGHYRGKASDTM--GRPCLPMSATVLOQTYHAHSDDLQLGLGKHNYCRNPDN 60
CC Db 30 CYTVNGADYRTQNTSLDGGKPCFLFWE--TFQHPYNTLKYPNKGGLGKHNYCRNPDG 87
CC
CC Qy 61 RRRPWCYV 68
CC Db 88 DVSPWCYI 95
CC
CC RESULT 24
CC KEMI_HUMAN
CC ID KEMI_HUMAN STANDARD; PRT; 475 AA.
CC AC Q96M08; Q9BY70; Q9UGS5; Q9UGU1;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
CC and the nose) (Dickkopf receptor).
CC GN KREMEN1 OR KREMEN.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN [1]

```

```

RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Nakamura T., Nakamura T.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kaneshori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Wakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kersey J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.P., Levensha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon I., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall N., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
RA Mcdermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,
RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkison P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tiliun Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
CC CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC CC to block Wnt/beta-catenin signaling (By similarity).
CC CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC CC -!- ALTERNATIVE PRODUCTS:
CC CC Event=Alternative splicing; Named isoforms=2;
CC CC Name=1;
CC CC IsoId=Q96M08-1; Sequence=Displayed;
CC CC Name=2;
CC CC IsoId=Q96M08-2; Sequence=VSP_003900;
CC CC Note=No experimental confirmation available;
CC CC -!- SIMILARITY: Contains 1 CUB domain.
CC CC -!- SIMILARITY: Contains 1 kringle domain.

```





CC use by non-profit institutions as long as its content is in no way

CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M62832; AAA41884.1; --  
 DR PIR; A40522; A40522.  
 DR HSP; P00747; 1PKK.  
 DR MEROPS; S01.233; --  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR003966; Prothrombin.  
 DR Pfam; PF00051; Kringle\_2.  
 DR PRINTS; PR00018; KRINGLE.  
 DR PRINTS; PR01505; PROTHROMBIN.  
 DR ProDom; PD000395; Kringle; 2.  
 DR SMART; SM00130; KR; 1.  
 DR PROSITE; PS00021; KRINGLE 1; 1.  
 DR PROSITE; PS00070; KRINGLE 2; 2.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; PARTIAL.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; PARTIAL.  
 DR PROSITE; PS00135; TRYPSIN\_SER; PARTIAL.  
 KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
 Tissue remodeling; Blood coagulation; Kringle; Repeat.  
 FT NON\_TER 1  
 FT DOMAIN <1 10 KRINGLE 3 (BY SIMILARITY).  
 FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).  
 FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).  
 FT DISULFID 34 112 BY SIMILARITY.  
 FT DISULFID 55 95 BY SIMILARITY.  
 FT DISULFID 83 107 BY SIMILARITY.  
 FT NON\_TER 169 169  
 SQ SEQUENCE 169 AA; 18401 MW; 77A54214C49D010C CRC64;  
 Query Match 29.3%; Score 149.5; DB 1; Length 169;  
 Best Local Similarity 37.8%; Pred. No. 2.9e-10;  
 Matches 34; Conservative 12; Mismatches 29; Indels 15; Gaps 6;  
 QY 3 CYEGNGHFRGKASTMTGRPCLPWNSATVLOQTHAHRSDALQL---GLGKHNYCRNPD 59  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 34 CYQGNGKSVRGTSSTNTTKKQSW-----VSNTPHSHKSTPANFPDSGL-EMNYCRNPD 87  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 60 N-RRRWYQVQGLKPLV--QECMVHDCAD 86  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 DB 88 NDORGWPCFT---TDFSVRWYCNLKRCS 114  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 RESULT 28  
 ID PLMN\_BOVIN STANDARD; PRT; 812 AA.  
 AC P06858; Q28162;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Plasminogen precursor (EC 3.4.21.7).  
 GN PLG.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Berglund L., Andersen M.D., Petersen T.E.;  
 ET "Cloning and characterization of the bovine plasminogen cDNA.";  
 RL Int. Dairy J. 5:593-603 (1995).  
 RN [2]  
 RN SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.  
 RP MEDLINE=85203906; PubMed=3846532;  
 RX Schaller J., Moser P.W., Danneberger-Muller G.A.K., Rosselet S.J.,  
 RA Kampfer U., Rickli E.E.;  
 ET "Complete amino acid sequence of bovine plasminogen. Comparison with human plasminogen.";  
 Eur. J. Biochem. 149:267-278 (1985).  
 [3]  
 RN SEQUENCE OF 706-812 FROM N.A.  
 RP MEDLINE=85023311; PubMed=6148961;  
 RX Malinowski D.P., Sadler J.E., Davie E.W.;  
 RA "Characterization of a complementary deoxyribonucleic acid coding for human and bovine plasminogen.";  
 RT Biochemistry 23:4243-4250 (1984).  
 RL [4]  
 RN CARBOHYDRATE-LINKAGE SITES.  
 RP MEDLINE=88185329; PubMed=3356193;  
 RX Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,  
 RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;  
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine plasminogen. Species specificity in relation to sialylation and fucosylation patterns.";  
 RL Eur. J. Biochem. 173:57-63 (1988).  
 CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION, AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN, LAMININ AND VON WILLEBRAND FACTOR.  
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa; higher selectivity than trypsin. Converts fibrin into soluble products.  
 CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN ACTIVATORS. BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.  
 CC -!- PTM: N-LINKED GLYCAN CONTAIN N-ACETYLGLUCOSAMINE AND SIALIC ACID. O-LINKED GLYCANS CONSIST OF GAL-GALNAc DISACCHARIDE WITH IS MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).  
 CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.  
 CC -!- SIMILARITY: Contains 5 kringle domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolyase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;  
KW SIGNAL.  
FT FT CHAIN 1 26  
FT FT CHAIN 27 812  
FT FT CHAIN 28 583  
FT FT CHAIN 584 812  
FT FT CHAIN 110 188  
FT FT DOMAIN 192 269  
FT FT DOMAIN 282 359  
FT FT DOMAIN 384 461  
FT FT DOMAIN 485 564  
FT FT DOMAIN 584 812  
FT FT CARBOHYD 315  
FT FT CARBOHYD 365  
FT FT ACT SITE 624  
FT FT ACT SITE 667  
FT FT ACT SITE 762  
FT FT ACT SITE 335  
FT FT CONFLICT 316  
FT FT CONFLICT 516  
FT FT CONFLICT 555  
FT FT CONFLICT 744  
FT FT CONFLICT 744  
SQ SEQUENCE 812 AA; 91216 MW; 38AA6A691E220946 CRC64;  
Query Match 29.2%; Score 149; DB 1; Length 812;  
Best Local Similarity 39.3%; Pred. No. 1.7e-09;  
Matches 35; Conservative 8; Mismatches 32; Indels 14; Gaps 5;  
QY 3 CYEGNGHYRGKASTDTGRCPLPNSATVLTQITVHAH---RSDALQLGKHNKNCNPD 59  
Db 384 CVHNGQSGYRGTSTTITGRKQSWSS-----MTPHRLKTPENYPNAGL-TMNYCRNPD 437  
QY 60 NERRPWCYVQGLKPLV--QECMWDHCDAD 86  
Db 438 ADKSPWCYT--TDFRVRWFCNKKCSE 463  
RESULT 29  
KRM2 HUMAN  
ID KRM2 HUMAN STANDARD; PRT: 462 AA.  
AC Q8NCW0; Q8N2J4; Q8NCW1; Q96GL8; Q9BTP9;  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Kremen protein 2 precursor (Kringle-containing protein marking the eye  
DE and the nose) (Dickkopf receptor 2).  
GN KREMEN2 OR KRM2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RA Tanaka S., Sugimachi K.;  
RT "Human Kremen2 and Wnt signaling.";  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).  
RC TISSUE=Ovarian carcinoma;  
RA Isozaki T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,  
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,  
RA Yamanoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,  
RA Masuhito Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,  
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;  
RT "NED0 human cDNA sequencing project.";  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).  
RC TISSUE=Brain, and Uterus;  
RX MEDLINE=22386257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carrincci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marita M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf  
CC to block Wnt/beta-catenin signaling. Forms a ternary complex with  
CC Dkk1 and LRP6 and induces rapid endocytosis and removal of the Wnt  
CC receptor LRP6 from the plasma membrane (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=1;  
CC IsoId=Q8NCW0-1; Sequence=Displayed;  
CC Name=2; Synonyms=Kremen2a;  
CC IsoId=Q8NCW0-2; Sequence=VSP\_050509, VSP\_050510;  
CC Name=3; Synonyms=Kremen2b;  
CC IsoId=Q8NCW0-3; Sequence=VSP\_050511, VSP\_050512;  
CC Name=4; Synonyms=Kremen2c;  
CC IsoId=Q8NCW0-4; Sequence=VSP\_050513, VSP\_050514;  
CC -!- SIMILARITY: Contains 1 CUB domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -!- SIMILARITY: Contains 1 WSC domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; AB086405; BAC00872.1; -  
CC EMBL; AB086355; BAC00823.1; -  
CC EMBL; AB086356; BAC00824.1; -  
CC EMBL; AB086357; BAC00825.1; -  
CC EMBL; AK027669; BAB55281.1; -  
CC EMBL; AK075033; BAC11365.1; -  
CC EMBL; BC003533; AAH03533.1; -  
CC EMBL; BC009383; AAH09383.1; -  
CC HSSP; P00750; 1PK2.  
CC Genew; HGNC:18797; KREMEN2.  
CC InterPro; IPR000859; CUB domain.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR002889; WSC.  
CC Pfam; PF00431; CUB; 1.  
CC Pfam; PF00051; kringle; 1.  
CC Pfam; PF01822; WSC; 1.  
CC PRINTS; PR00018; KRINGLE.  
CC ProDom; PD000395; Kringle; 1.  
CC SMART; SM00042; CUB; 1.  
CC SMART; SM00130; KR; 1.  
CC SMART; SM00321; WSC; 1.  
CC PROSITE; PS01180; CUB; 1.  
CC PROSITE; PS00021; KRINGLE\_2; 1.  
CC PROSITE; PS00070; KRINGLE\_2; 1.  
KW Wnt signaling pathway; Glycoprotein; Kringle; Signal; Transmembrane;  
KW Alternative splicing.



Best Local Similarity 41.5%; Pred. No. 1.9e-09;  
 Matches 27; Conservative 7; Mismatches 28; Indels 3; Gaps 1;  
 Qy 3 CYEGNGHYRGKASTDTMGKPCLPWNSATVLCQTYHAHRSDALQLGLGKHNYCRNPDNR 62  
 Db 379 CYQNGQTYRGTSSTTITGKCKQPWTSRPHRHSKTPENYPADLTM--NYCRNPDGDK 435  
 Qy 63 RPWCY 67  
 Db 436 GPWCY 440

Search completed: December 3, 2003, 14:40:04  
 Job time : 3.82759 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:02 ; Search time 13.8621 Seconds  
(without alignments)  
1007.637 Million cell updates/sec

Title: US-09-880-503-1

Perfect score: 510

Sequence: 1 KTCYEGNGHFYRGKASTDTM.....QVGLKPLVQRCMVHDCADGK 88

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_15Jun03:\*  
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	100.0	88	AAE16542	Human urokinase-ty
2	510	100.0	96	AAE16550	Human uPA, kringle
3	510	100.0	135	AAE16545	Human urokinase-ty
4	510	100.0	143	AAE16549	Human uPA amino te
5	510	100.0	337	AAAG75492	Human colon cancer
6	510	100.0	337	ABP41795	Human ovarian anti
7	510	100.0	365	ABP68854	Delta 1-46 urokina
8	510	100.0	378	AAW13635	Human prourokinase
9	510	100.0	386	AAE66266	Bifunctional uroki

10	510	100.0	389	11	AAW13636	Human prourokinase
11	510	100.0	390	16	AAE66245	Bifunctional uroki
12	510	100.0	390	16	AAE66247	Bifunctional uroki
13	510	100.0	392	16	AAE66254	Bifunctional uroki
14	510	100.0	392	16	AAE66255	Bifunctional uroki
15	510	100.0	392	16	AAE66256	Bifunctional uroki
16	510	100.0	392	16	AAE66257	Bifunctional uroki
17	510	100.0	392	16	AAE66258	Bifunctional uroki
18	510	100.0	392	16	AAE66259	Bifunctional uroki
19	510	100.0	392	16	AAE66260	Bifunctional uroki
20	510	100.0	392	16	AAE66261	Bifunctional uroki
21	510	100.0	392	16	AAE66263	Bifunctional uroki
22	510	100.0	393	16	AAE66264	Bifunctional uroki
23	510	100.0	393	16	AAE66244	Bifunctional uroki
24	510	100.0	393	16	AAE66249	Bifunctional uroki
25	510	100.0	393	16	AAE66250	Bifunctional uroki
26	510	100.0	393	16	AAE66251	Bifunctional uroki
27	510	100.0	393	16	AAE66252	Bifunctional uroki
28	510	100.0	393	16	AAE66253	Bifunctional uroki
29	510	100.0	393	17	AAE99885	M36: fibrinolytic
30	510	100.0	393	17	AAE99596	Chimeric protein M
31	510	100.0	393	17	AAE99597	Chimeric protein M
32	510	100.0	395	15	AAE47902	Pro-urokinase deri
33	510	100.0	395	16	AAE66262	Bifunctional uroki
34	510	100.0	395	16	AAE66265	Bifunctional uroki
35	510	100.0	396	16	AAE66246	Bifunctional uroki
36	510	100.0	397	16	AAE66248	Bifunctional uroki
37	510	100.0	401	11	AAW13637	Human prourokinase
38	510	100.0	403	23	AAE16547	Human urokinase-ty
39	510	100.0	410	13	AAE23794	Prourokinase mutan
40	510	100.0	410	13	AAE23795	Prourokinase doub
41	510	100.0	411	6	AAE50871	Sequence encoded b
42	510	100.0	411	11	AAW13634	Human native prou
43	510	100.0	411	11	AAE05117	UK-53 as encoded b
44	510	100.0	411	11	AAE06244	Urokinase precurs
45	510	100.0	411	11	AAE07902	Human pro-urokinas

#### ALIGNMENTS

#### RESULT 1

AAE16542  
ID AAE16542 standard; Protein; 88 AA.

XX AAE16542;

AC AAE16542;

DT 09-APR-2002 (first entry)

XX Human urokinase-type plasminogen activator (uPA) kringle.

XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
clotting disorder; uterine contraction disorder; respiratory disease;  
male impotence; adult respiratory distress syndrome.

OS Homo sapiens.

XX WO200197752-A2.

PN 27-DEC-2001.

PD 13-JUN-2001; 2001WO-US18976.

XX 20-JUN-2000; 2000US-212874P.

PF (UYPE-) UNIV PENNSYLVANIA.

XX Cines DB, Higazi AA;

XX WPI; 2002-122240/16.

DR N-PSDB; AAD27075.  
 XX Composition for modulating muscle cell and tissue contractility for  
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
 PT comprising domains from urokinase-type plasminogen activator -  
 XX  
 XX Claim 1; Fig 1A; 117pp; English.  
 XX The invention relates to a composition comprising one or more domains of  
 CC urokinase-type plasminogen activator (uPA). The composition is used to  
 CC modulate the contractility and angiogenic activity of a mammalian muscle,  
 CC endothelial cell or tissue. The composition is used for treating stroke,  
 CC hypertension, atherosclerosis, heart attack, microvascular  
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
 CC disorders, tumour cell angiogenesis, pulmonary fibrosis, asthma, tumour cell  
 CC invasion, tumour cell angiogenesis, pulmonary fibrosis, asthma, tumour cell  
 CC contraction disorder, wound healing, clotting disorder, uterine  
 CC diabetic retinopathy, wound healing, clotting disorder, uterine  
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
 CC contraction disorder, male impotence, respiratory disease or condition  
 CC associated with chronic intrapulmonary fibrin formation, and a disorder  
 CC hypertension, microvascular thrombotic occlusion, and a disorder  
 CC associated with chronic intrapulmonary fibrin formation. The present  
 CC sequence is human urokinase-type plasminogen activator (uPA) kringle.  
 XX  
 XX Sequence 88 AA;  
 SQ  
 Query Match 100.0%; Score 510; DB 23; Length 88;  
 Best Local Similarity 100.0%; Pred. NO. 1.1e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60  
 Db 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60  
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 Db 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 RESULT 2  
 AAE16550  
 ID AAE16550 standard; Protein; 96 AA.  
 AC AAE16550;  
 XX  
 XX 09-APR-2002 (first entry)  
 DT  
 DE Human uPA kringle and connecting peptide.  
 KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
 KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
 KW clotting disorder; uterine contraction disorder; respiratory disease;  
 KW adult respiratory distress syndrome; male impotence.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200197752-A2.  
 PN  
 XX 27-DEC-2001.  
 PD  
 XX 13-JUN-2001; 2001WO-US18976.  
 PF  
 XX 20-JUN-2000; 2000US-212874P.  
 PR  
 XX (UYPE-) UNIV PENNSYLVANIA.  
 PA  
 XX Cines DB, Higazi AA;  
 PI  
 XX WPI; 2002-122240/16.  
 DR  
 XX N-PSDB; AAD27083.  
 XX Composition for modulating muscle cell and tissue contractility for

PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
 PT comprising domains from urokinase-type plasminogen activator -  
 XX Claim 25; Fig 1I; 117pp; English.  
 XX The invention relates to a composition comprising one or more domains of  
 CC urokinase-type plasminogen activator (uPA). The composition is used to  
 CC modulate the contractility and angiogenic activity of a mammalian muscle,  
 CC endothelial cell or tissue. The composition is used for treating stroke,  
 CC hypertension, atherosclerosis, heart attack, microvascular  
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
 CC disorders, tumour cell angiogenesis, pulmonary fibrosis, asthma, tumour cell  
 CC invasion, tumour cell angiogenesis, pulmonary fibrosis, asthma, tumour cell  
 CC contraction disorder, wound healing, clotting disorder, uterine  
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
 CC contraction disorder, male impotence, respiratory disease or condition  
 CC associated with chronic intrapulmonary fibrin formation. The present  
 CC sequence is human urokinase-type plasminogen activator (uPA) kringle  
 CC and connecting peptide.  
 XX  
 XX Sequence 96 AA;  
 SQ  
 Query Match 100.0%; Score 510; DB 23; Length 96;  
 Best Local Similarity 100.0%; Pred. NO. 1.2e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60  
 Db 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60  
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 Db 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 RESULT 3  
 AAE16545  
 ID AAE16545 standard; Protein; 135 AA.  
 AC AAE16545;  
 XX  
 XX 09-APR-2002 (first entry)  
 DT  
 DE Human urokinase-type plasminogen activator amino terminal fragment (ATF).  
 KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
 KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
 KW clotting disorder; uterine contraction disorder; respiratory disease;  
 KW adult respiratory distress syndrome; amino terminal fragment; ATF;  
 KW male impotence.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200197752-A2.  
 PN  
 XX 27-DEC-2001.  
 PD  
 XX 13-JUN-2001; 2001WO-US18976.  
 PF  
 XX 20-JUN-2000; 2000US-212874P.  
 PR  
 XX (UYPE-) UNIV PENNSYLVANIA.  
 PA  
 XX Cines DB, Higazi AA;  
 PI  
 XX WPI; 2002-122240/16.  
 DR  
 XX N-PSDB; AAD27078.  
 XX Composition for modulating muscle cell and tissue contractility for  
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,

PT comprising domains from urokinase-type plasminogen activator -  
XX  
PS Claim 11; Fig 1D; 117pp; English.

XX The invention relates to a composition comprising one or more domains of  
CC urokinase-type plasminogen activator (uPA). The composition is used to  
CC modulate the contractility and angiogenic activity of a mammalian muscle,  
CC endothelial cell or tissue. The composition is used for treating stroke,  
CC hypotension, hypertension, atherosclerosis, heart attack, microvascular  
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
CC diabetic retinopathy, wound healing, clotting disorder, uterine  
CC contraction disorder, male impotence, respiratory disease or condition  
CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
CC hypertension, microvascular thrombotic occlusion, and a disorder  
CC associated with chronic intrapulmonary fibrin formation. The present  
CC sequence is human urokinase-type plasminogen activator (uPA) amino  
CC terminal fragment (ATF).  
XX

SQ Sequence 135 AA;

Query Match 100.0%; Score 510; DB 23; Length 135;  
Best Local Similarity 100.0%; Pred. No. 1.7e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
DB 48 KTCYEGNGHFYRGKASTDTWGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 107  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 4

AAE16549  
ID AAE16549 standard; Protein; 143 AA.

AC AAE16549;

DT 09-APR-2002 (first entry)

XX Human uPA amino terminal fragment (ATF) and connecting peptide.

DE Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
KW clotting disorder; uterine contraction disorder; respiratory disease;  
KW adult respiratory distress syndrome; amino terminal fragment; ATF;  
KW male impotence.

XX Homo sapiens.

OS  
XX WO200197752-A2.

PN 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US18976.

XX 20-JUN-2000; 2000US-212874P.

XX (UYPE-) UNIV PENNSYLVANIA.

PA Cines DB, Higazi AA;

XX WPI; 2002-122240/16.

DR N-PSDB; AAD27082.

XX Composition for modulating muscle cell and tissue contractility for  
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
PT comprising domains from urokinase-type plasminogen activator -

XX  
PS  
XX

Claim 24; Fig 1H; 117pp; English.

CC The invention relates to a composition comprising one or more domains of  
CC urokinase-type plasminogen activator (uPA). The composition is used to  
CC modulate the contractility and angiogenic activity of a mammalian muscle,  
CC endothelial cell or tissue. The composition is used for treating stroke,  
CC hypotension, hypertension, atherosclerosis, heart attack, microvascular  
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
CC diabetic retinopathy, wound healing, clotting disorder, uterine  
CC contraction disorder, male impotence, respiratory disease or condition  
CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
CC hypertension, microvascular thrombotic occlusion, and a disorder  
CC associated with chronic intrapulmonary fibrin formation. The present  
CC sequence is human urokinase-type plasminogen activator (uPA) amino  
CC terminal fragment (ATF) and connecting peptide.  
XX

SQ Sequence 143 AA;

Query Match 100.0%; Score 510; DB 23; Length 143;  
Best Local Similarity 100.0%; Pred. No. 1.8e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
DB 48 KTCYEGNGHFYRGKASTDTWGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 107  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 5

AAAG75492  
ID AAG75492 standard; Protein; 337 AA.

AC AAG75492;

DT 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:6256.

DE Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma; chromosome 10.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

DR N-PSDB; AAH34897.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers -

PS Claim 11; Page 7707-7708; 9803pp; English.

XX AAH32943 to AAH37195 and AAG773514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon

CC cancer antigens have cytostatic activity and can be used in gene  
 CC therapy and vaccine production. N and P may be used in the prevention,  
 CC diagnosis and treatment of diseases associated with inappropriate P  
 CC expression. For example, N and P may be used to treat disorders  
 CC associated with decreased expression by rectifying mutations or deletions  
 CC in a patient's genome that affect the activity of P by expressing  
 CC inactive proteins or to supplement the patients own production of P.  
 CC Additionally, N may be used to produce the colon cancer-associated Ps,  
 CC by inserting the nucleic acids into a host cell and culturing the cell  
 CC to express the proteins. N and P can be used in the prevention, diagnosis  
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
 CC and AAH77789 represent sequences used in the exemplification of the  
 CC present invention.  
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
 CC missing at time of publication, meaning no sequences are present for  
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX SQ Sequence 337 AA;

Query Match 100.0%; Score 510; DB 22; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 4e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60

DB 74 KTCYEGNGHFFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 133

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88

DB 134 RRRPWCYVQVGLKPLVQECMVHDCADGK 161

RESULT 6

ABP41795

ID ABP41795 standard; Protein; 337 AA.

XX AC ABP41795;

DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HVVCB79, SEQ ID NO:2927.

XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW anti-inflammatory; gynaecological; reproductive; chromosome 10q24.

XX OS Homo sapiens.

XX FN WO200200677-A1.

XX PD 03-JAN-2002.

XX FF 07-JUN-2001; 2001WO-US18569.

XX XX 07-JUN-2000; 2000US-209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX XX WPI; 2002-147878/19.

DR N-PSDB; ABQ54872.

XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.  
 PT ovarian cancer), immune disorders, cardiovascular disorders and

PT neurological diseases -

XX PS Claim 11; SEQ ID No 2927; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 337 AA;

Query Match 100.0%; Score 510; DB 23; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 4e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60

DB 74 KTCYEGNGHFFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 133

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88

DB 134 RRRPWCYVQVGLKPLVQECMVHDCADGK 161

RESULT 7

AAH68854

ID AAH68854 standard; protein; 365 AA.

XX AC AAH68854;

XX AC

DT 25-MAR-2003 (updated)

DT 22-NOV-1995 (first entry)

XX DE Delta 1-46 urokinase.

XX KW Human; des-epidermal growth factor homologous plasminogen activator;

XX KW uPA; liver membrane; reduced affinity; BGF homologous; thrombosis;

XX KW thrombolytic; increased half-life; urokinase.

XX OS Homo sapiens (engineered).

XX XX Key Location/Qualifiers

FT FT Misc-difference 1

FT FT /note= "amino acids 1-46 of wild-type urokinase  
 have been deleted"

XX US5376547-A.

PD XX 27-DEC-1994.  
 PF XX 29-JAN-1988; 88US-0150267.  
 PR XX 30-JAN-1987; 87US-0008795.  
 PR XX 29-JAN-1988; 88US-0150267.  
 XX XX (AMHP ) AMERICAN HOME PROD CORP.  
 XX XX Hung PP, Kalyan NK, Lee SL;  
 XX XX WPI; 1995-043464/06.  
 DR XX New modified plasminogen activator cpds. - having regions removed  
 PT XX to reduce affinity for liver membranes and increase circulation  
 PT XX half-life.  
 XX XX Claim 1; Page ?; 26pp; English.  
 PS XX Amino acid residues 1-46 contain the EGF region of human urokinase.  
 CC XX Deletion of this region results in a plasminogen activator with  
 CC XX reduced affinity for liver cell membranes; the mutant protein is  
 CC XX not cleared from the circulation as rapidly as is wild-type tPA.  
 CC XX The specification only gives the sequence around the deletion and  
 CC XX not the full-length sequence of "delta 1-46 urokinase"; the  
 CC XX sequence in AAR68854 has been obtained by amending a previously  
 CC XX disclosed wild-type human urokinase sequence (from W09501427)  
 CC XX according to the description given in Example 3.  
 CC XX (Updated on 25-MAR-2003 to correct PF field.)  
 XX XX  
 SQ Sequence 365 AA;  
 Query Match 100.0%; Score 510; DB 16; Length 365;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60  
 DB 2 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 61  
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
 RESULT 8  
 AAW13635  
 ID AAW13635 standard; Protein; 378 AA.  
 XX AC AAW13635;  
 XX DT 04-JUN-1997 (first entry)  
 XX DE Human prourokinase variant lacking entire EGF domain.  
 XX KW Human; prourokinase; hPUK; variant; half-life; increase; EGF;  
 KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Region 1..9  
 FT /note= "residues 1-9 of native hPUK"  
 FT Region 10..378  
 FT /note= "residues 43-411 of native hPUK"  
 FT Misc-difference 144  
 FT /note= "corresponds to TAC codon"  
 XX EP398361-A.  
 PN 22-NOV-1990.  
 PD XX

PF 18-MAY-1990; 90EP-0109472.  
 XX 22-FEB-1990; 90JP-0042020.  
 PR 18-MAY-1989; 89JP-0126433.  
 PR 03-JUL-1986; 86JP-0156936.  
 PR 18-FEB-1987; 87JP-0036495.  
 PR 18-MAY-1989; 89JP-0126434.  
 XX (GREC ) GREEN CROSS CORP.  
 XX Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;  
 PI Morita M, Tanabe T;  
 XX WPI; 1990-350146/47.  
 DR N-PSDB; AAT61672.  
 XX Human pro-urokinase variants - deficient in loop regions of  
 PT epidermal growth factor, showing long blood half-life, as  
 PT fibrinolytic agent  
 XX Claim 1; Page -; 22pp; English.  
 PS New variants of human prourokinase (hPUK) comprise a hPUK deficient  
 CC in (i) at least part of the first loop region of the epidermal growth  
 CC factor (EGF) domain; (ii) at least part of the first loop and at  
 CC least part of the second loop; or (iii) at least part of the third  
 CC loop. The hPUK variants show an increased blood half-life comparable  
 CC to that of the whole EGF domain-deficient hPUK variant and urokinase  
 CC while retaining the same properties as those of hPUK. They have  
 CC potent thrombolytic activity and very little tendency to cause  
 CC spontaneous bleeding. The present sequence represents a specific variant  
 CC of hPUK which lacks the entire EGF domain; the sequence does not  
 CC appear in the specification and has been created using the  
 CC wild-type hPUK sequence and the junction sequence after  
 CC deletion, both of which are given (in Fig 1 and in Fig 2(3),  
 CC respectively).  
 XX SQ Sequence 378 AA;  
 Query Match 100.0%; Score 510; DB 11; Length 378;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60  
 DB 15 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 74  
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 DB 75 RRRPWCYVQVGLKPLVQECMVHDCADGK 102  
 RESULT 9  
 AAR66266  
 ID AAR66266 standard; protein; 386 AA.  
 XX AC AAR66266;  
 XX DT 25-MAR-2003 (updated)  
 DT 17-AUG-1995 (first entry)  
 XX DE Bifunctional urokinase variant M33.  
 XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Region 1..365  
 FT /label= M4  
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
 FT Disulfide-bond 4..85

FT Disulfide-bond 25...67  
 FT Disulfide-bond 56...80  
 FT Disulfide-bond 102...233  
 FT Disulfide-bond 143...159  
 FT Disulfide-bond 151...222  
 FT Disulfide-bond 247...316  
 FT Disulfide-bond 279...295  
 FT Disulfide-bond 306...334  
 FT Region 365...366  
 FT /label= X1  
 FT /note= "peptide bond"  
 FT /note= 366...386  
 FT /label= Y1  
 PN DE4323754-Cl.  
 PD 01-DEC-1994.  
 XX 15-JUL-1993; 93DE-4323754.  
 XX 15-JUL-1993; 93DE-4323754.  
 XX (CHEF ) GRUENTHAL GMBH.  
 XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;  
 PI Wndt S;  
 XX WPI; 1995-015191/03.  
 XX New bifunctional urokinase derivs and related plasmids - with  
 PT improved fibrinolytic and thrombin inhibiting activities, for  
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc  
 XX Example 1; Page 11 and Fig 1; 34pp; German.  
 PS Bifunctional urokinase derivatives corresponding to the formula  
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266  
 CC are specific examples of such derivs. which have both improved  
 CC fibrinolytic and thrombin-inhibiting activities, compared to known  
 CC plasminogen activators or thrombin inhibitors. The proteins are  
 CC useful as thrombolytic agents, e.g. for treatment of arterial  
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction  
 CC and pulmonary embolism.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 386 AA;  
 Query Match 100.0%; Score 510; DB 16; Length 386;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60  
 Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 61  
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
 RESULT 10  
 ID AAW13636  
 XX AAW13636 standard; Protein; 389 AA.  
 AC AAW13636;  
 XX  
 DT 04-JUN-1997 (first entry)  
 XX Human prourokinase variant lacking EGF domain loops 1 and 2.  
 DE Human; prourokinase; hPUK; variant; half-life; increase; EGF;  
 KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.  
 XX

OS Homo sapiens.  
 OS Synthetic.  
 XX Key Location/Qualifiers  
 FT Region 1..10  
 FT /note= "residues 1-10 of native hPUK"  
 FT Region 11..389  
 FT /note= "residues 33-411 of native hPUK"  
 FT Misc-difference 155  
 FT /note= "corresponds to TAC codon"  
 XX  
 PN EP398361-A.  
 XX  
 PD 22-NOV-1990.  
 XX 18-MAY-1990; 90EP-0109472.  
 XX 22-FEB-1990; 90JP-0042020.  
 PR 18-MAY-1989; 89JP-0126433.  
 PR 03-JUL-1986; 86JP-0156936.  
 PR 18-FEB-1987; 87JP-0036495.  
 PR 18-MAY-1989; 89JP-0126434.  
 XX  
 PA (GREG ) GREEN CROSS CORP.  
 XX Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;  
 PI Morita M, Tanabe T;  
 XX  
 DR WPI; 1990-350146/47.  
 DR N-PSDB; AAT61673.  
 XX Human pro-urokinase variants - deficient in loop regions of  
 PT epidermal growth factor, showing long blood half-life, as  
 PT fibrinolytic agent  
 XX  
 PS Claim 6; Page -; 22pp; English.  
 XX New variants of human prourokinase (hPUK) comprise a hPUK deficient  
 CC in (i) at least part of the first loop region of the epidermal growth  
 CC factor (EGF) domain; (ii) at least part of the first loop and at  
 CC least part of the second loop; or (iii) at least part of the third  
 CC loop. The hPUK variants show an increased blood half-life comparable  
 CC to that of the whole EGF domain-deficient hPUK variant and urokinase  
 CC while retaining the same properties as those of hPUK. They have  
 CC potent thrombolytic activity and very little tendency to cause  
 CC spontaneous bleeding. The present sequence represents a specific  
 CC variant of hPUK which lacks loops 1 and 2 of the EGF domain; the  
 CC sequence does not appear in the specification and has been created  
 CC using the wild-type hPUK sequence and the junction sequence after  
 CC deletion, both of which are given (in Fig 1 and on page 8,  
 CC respectively).  
 SQ Sequence 389 AA;  
 Query Match 100.0%; Score 510; DB 11; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60  
 Db 26 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 85  
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 Db 86 RRRPWCYVQVGLKPLVQECMVHDCADGK 113  
 RESULT 11  
 ID AAR66245  
 XX AAR66245 standard; protein; 390 AA.  
 AC AAR66245;  
 XX

```
DT 25-MAR-2003 (updated)
DE 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M12.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT /label= X1
FT /label= Y1
XX
PN DB4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heintel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Whendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 390 AA;
Query Match 100.0%; Score 510; DB 16; Length 390;
Best Local Similarity 100.0%; Pred. No. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQOYTHAHRSDALQLGLGKHNCRPN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQOYTHAHRSDALQLGLGKHNCRPN 61
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGK 88
Db 62 RRRPWCYVQGLKPLVQECMVHDCADGK 89
```

```
RESULT 12
AAR66247
ID AAR66247 standard; protein; 390 AA.
XX
XX AAR66247;
XX
XX 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M14.
DE
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT /label= X1
FT /label= Y1
XX
XX DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heintel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Whendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 390 AA;
Query Match 100.0%; Score 510; DB 16; Length 390;
Best Local Similarity 100.0%; Pred. No. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQOYTHAHRSDALQLGLGKHNCRPN 60
```

4 KTCYEGNGHFGKASTDTWGRPCLPWNATVLTQTYHAHRSDALQGLGKKNYCRNPDN 61

QY 61 RRRPWCYVQGLKPLVOCNMVHDCADGK 88  
Db 62 RRRPWCYVQGLKPLVOCNMVHDCADGK 89

RESULT 13  
AAR66254  
ID AAR66254 standard; protein; 392 AA.  
XX AAR66254;  
XX 25-MAR-2003 (updated)  
DT 17-AUG-1995 (first entry)  
XX Bifunctional urokinase variant M21.  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
KW urokinase; variant; mutein.  
XX Synthetic.  
XX Key  
FH Location/Qualifiers  
FT 1..365  
FT /label= M4  
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"  
FT Disulfide-bond 4..82  
FT Disulfide-bond 25..87  
FT Disulfide-bond 56..80  
FT Disulfide-bond 102..233  
FT Disulfide-bond 143..159  
FT Disulfide-bond 151..222  
FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Disulfide-bond 366..371  
FT Region /label= X1  
FT Region /label= Y1  
FT Region /label= Y1  
PN DE4323754-C1.  
XX 01-DEC-1994.  
XX 15-JUL-1993; 93DE-4323754.  
XX 15-JUL-1993; 93DE-4323754.  
XX (CHEF ) GRUENENTHAL GMBH.  
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;  
PI Wendt S;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with  
PT improved fibrinolytic and thrombin inhibiting activities, for  
X treating cardiac and cerebral infarct, pulmonary embolism, etc  
S Example 1; Page 10 and Fig 1; 34pp; German.  
X Bifunctional urokinase derivatives corresponding to the formula  
X M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266  
X are specific examples of such derivs which have both improved  
X fibrinolytic and thrombin-inhibiting activities compared to known  
X plasminogen activators or thrombin inhibitors. The proteins are  
X useful as thrombolytic agents, e.g. for treatment of arterial  
X occlusions, deep vein thrombosis, cardiac and cerebral infarction  
X and pulmonary embolism.  
X Updated on 25-MAR-2003 to correct PN field.)

QY 1 KTCYEGNGHFGKASTDTWGRPCLPWNATVLTQTYHAHRSDALQGLGKKNYCRNPDN 60  
Db 2 KTCYEGNGHFGKASTDTWGRPCLPWNATVLTQTYHAHRSDALQGLGKKNYCRNPDN 61  
QY 61 RRRPWCYVQGLKPLVOCNMVHDCADGK 88  
Db 62 RRRPWCYVQGLKPLVOCNMVHDCADGK 89  
RESULT 14  
AAR66255  
ID AAR66255 standard; protein; 392 AA.  
XX AAR66255;  
XX 25-MAR-2003 (updated)  
DT 17-AUG-1995 (first entry)  
XX Bifunctional urokinase variant M22.  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
KW urokinase; variant; mutein.  
XX Synthetic.  
XX Key  
FH Location/Qualifiers  
FT 1..365  
FT /label= M4  
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"  
FT Disulfide-bond 4..82  
FT Disulfide-bond 25..87  
FT Disulfide-bond 56..80  
FT Disulfide-bond 102..233  
FT Disulfide-bond 143..159  
FT Disulfide-bond 151..222  
FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Disulfide-bond 366..371  
FT Region /label= X1  
FT Region /label= Y1  
FT Region /label= Y1  
PN DE4323754-C1.  
XX 01-DEC-1994.  
XX 15-JUL-1993; 93DE-4323754.  
XX 15-JUL-1993; 93DE-4323754.  
XX (CHEF ) GRUENENTHAL GMBH.  
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;  
PI Wendt S;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with  
PT improved fibrinolytic and thrombin inhibiting activities, for  
X treating cardiac and cerebral infarct, pulmonary embolism, etc  
S Example 1; Page 10 and Fig 1; 34pp; German.  
X Bifunctional urokinase derivatives corresponding to the formula  
X M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266  
X are specific examples of such derivs which have both improved  
X fibrinolytic and thrombin-inhibiting activities compared to known  
X plasminogen activators or thrombin inhibitors. The proteins are  
X useful as thrombolytic agents, e.g. for treatment of arterial  
X occlusions, deep vein thrombosis, cardiac and cerebral infarction  
X and pulmonary embolism.  
X Updated on 25-MAR-2003 to correct PN field.)

CC



```
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 392 AA;
Query Match 100.0%; Score 510; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWN SATVLQQT YHAHRS DALQLGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTWGRPCLPWN SATVLQQT YHAHRS DALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
RESULT 15
AAR66256
ID AAR66256 standard; protein; 392 AA.
XX
AC AAR66256;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M23.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365 /label= M4
FT Disulfide-bond 4..85 /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region 372..392 /label= Y1
XX
DE4323754-Cl.
XX
PD 01-DEC-1994.
XX
PF 15-JUL-1993; 93DE-4323754.
XX
PR 15-JUL-1993; 93DE-4323754.
XX
PA (CHEF ) GRUENENTHAL GMBH.
XX
PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX
DR WPI; 1995-015191/03.
XX
PT New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
```

```
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
CC Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 392 AA;
Query Match 100.0%; Score 510; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWN SATVLQQT YHAHRS DALQLGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTWGRPCLPWN SATVLQQT YHAHRS DALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
RESULT 16
AAR66257
ID AAR66257 standard; protein; 392 AA.
XX
AC AAR66257;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M24.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365 /label= M4
FT Disulfide-bond 4..85 /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region 372..392 /label= Y1
XX
DE4323754-Cl.
XX
PD 01-DEC-1994.
XX
PF 15-JUL-1993; 93DE-4323754.
XX
PR 15-JUL-1993; 93DE-4323754.
XX
PA (CHEF ) GRUENENTHAL GMBH.
XX
```

UR WPI: 1995-015191/03.  
 XX New bifunctional urokinase derivs and related plasmids - with  
 PT improved fibrinolytic and thrombin inhibiting activities, for  
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc  
 XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;  
 PS Wnendt S;  
 XX WPI: 1995-015191/03.  
 XX Example 1; Page 11 and Fig 1; 34pp; German.  
 XX Bifunctional urokinase derivatives corresponding to the formula  
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266  
 CC are specific examples of such derivs. which have both improved  
 CC fibrinolytic and thrombin-inhibiting activities, compared to known  
 CC plasminogen activators or thrombin inhibitors. The proteins are  
 CC useful as thrombolytic agents, e.g. for treatment of arterial  
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction  
 CC and pulmonary embolism.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX Sequence 392 AA;  
 SQ Query Match 100.0%; Score 510; DB 16; Length 392;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPNNSATVLQOTYHAHRSALQLGLGKHNYCRNPDN 60  
 DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPNNSATVLQOTYHAHRSALQLGLGKHNYCRNPDN 61  
 QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
 DB 62 RRRPWCYVQVGLKPLVQECWVHDCADGK 89  
 RESULT 17  
 ID AAR66258  
 AC AAR66258;  
 XX 25-MAR-2003 (updated)  
 DT 17-AUG-1995 (first entry)  
 XX Bifunctional urokinase variant M25.  
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 XX Synthetic.  
 PH Key Location/Qualifiers  
 FT Region 1..365 /label= M4  
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"  
 FT Disulfide-bond 4..85  
 FT Disulfide-bond 25..67  
 FT Disulfide-bond 56..80  
 FT Disulfide-bond 102..233  
 FT Disulfide-bond 143..159  
 FT Disulfide-bond 151..222  
 FT Disulfide-bond 247..316  
 FT Disulfide-bond 279..295  
 FT Disulfide-bond 306..334  
 FT Region 366..371 /label= X1  
 FT 372..392 /label= Y1  
 DE4323754-Cl.  
 -DEC-1994.

Saunders DJ, Schneider J, Steffens GJ;  
 WPI: 1995-015191/03.  
 New bifunctional urokinase derivs and related plasmids - with  
 improved fibrinolytic and thrombin inhibiting activities, for  
 treating cardiac and cerebral infarct, pulmonary embolism, etc  
 Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;  
 Wnendt S;  
 WPI: 1995-015191/03.  
 Example 1; Page 11 and Fig 1; 34pp; German.  
 Bifunctional urokinase derivatives corresponding to the formula  
 M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266  
 are specific examples of such derivs. which have both improved  
 fibrinolytic and thrombin-inhibiting activities, compared to known  
 plasminogen activators or thrombin inhibitors. The proteins are  
 useful as thrombolytic agents, e.g. for treatment of arterial  
 occlusions, deep vein thrombosis, cardiac and cerebral infarction  
 and pulmonary embolism.  
 (Updated on 25-MAR-2003 to correct PN field.)  
 Sequence 392 AA;  
 Query Match 100.0%; Score 510; DB 16; Length 392;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPNNSATVLQOTYHAHRSALQLGLGKHNYCRNPDN 60  
 DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPNNSATVLQOTYHAHRSALQLGLGKHNYCRNPDN 61  
 QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
 DB 62 RRRPWCYVQVGLKPLVQECWVHDCADGK 89  
 RESULT 18  
 ID AAR66259  
 AC AAR66259 standard; protein; 392 AA.  
 XX 25-MAR-2003 (updated)  
 DT 17-AUG-1995 (first entry)  
 XX Bifunctional urokinase variant M26.  
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 XX Synthetic.  
 PH Key Location/Qualifiers  
 FT Region 1..365 /label= M4  
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"  
 FT Disulfide-bond 4..85  
 FT Disulfide-bond 25..67  
 FT Disulfide-bond 56..80  
 FT Disulfide-bond 102..233  
 FT Disulfide-bond 143..159  
 FT Disulfide-bond 151..222  
 FT Disulfide-bond 247..316  
 FT Disulfide-bond 279..295  
 FT Disulfide-bond 306..334  
 FT Region 366..371 /label= X1  
 FT 372..392 /label= Y1  
 DE4323754-Cl.  
 -DEC-1994.

```
FT FT /label= X1
FT FT 372..392
FT FT /label= Y1
XX XX
PN DE4323754-Cl.
XX XX
XX 01-DEC-1994.
XX XX
XX 15-JUL-1993; 93DE-4323754.
XX XX
XX 15-JUL-1993; 93DE-4323754.
XX XX
XX (CHEF ) GRUENENTHAL GMBH.
XX XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Whendt S;
XX XX
XX WPI; 1995-015191/03.
XX XX
XX Bifunctional urokinase derivs and related plasmids - with
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX XX
XX Sequence 392 AA;
XX XX
XX Query Match 100.0%; Score 510; DB 16; Length 392;
XX Best Local Similarity 100.0%; Pred. No. 4.6e-41;
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
RESULT 19
AAR66260
ID AAR66260 standard; protein; 392 AA.
XX XX
XX AAR66260;
XX XX
XX 25-MAR-2003 (updated)
XX 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M27.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.
XX Synthetic.
XX Key Location/Qualifiers
XX Region 1..365
XX /label= M4
XX /notes= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
```

```
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..371
FT /label= X1
FT Region 372..392
FT /label= Y1
XX XX
XX DE4323754-Cl.
XX XX
XX 01-DEC-1994.
XX XX
XX 15-JUL-1993; 93DE-4323754.
XX XX
XX 15-JUL-1993; 93DE-4323754.
XX XX
XX (CHEF ) GRUENENTHAL GMBH.
XX XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Whendt S;
XX XX
XX WPI; 1995-015191/03.
XX XX
XX Bifunctional urokinase derivs and related plasmids - with
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX XX
XX Sequence 392 AA;
XX XX
XX Query Match 100.0%; Score 510; DB 16; Length 392;
XX Best Local Similarity 100.0%; Pred. No. 4.6e-41;
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
RESULT 20
AAR66261
ID AAR66261 standard; protein; 392 AA.
XX XX
XX AAR66261;
XX XX
XX 25-MAR-2003 (updated)
XX 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M28.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.
XX Synthetic.
XX XX
```

RESULT 22  
AAR66264

```
ID AAR66264 standard; protein; 392 AA.
XX AC AAR66264;
XX
XX 25-MAR-2003 (updated)
XX DT 17-AUG-1995 (first entry)
XX DE Bifunctional urokinase variant M31.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX KW urokinase; variant; mutein.
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX FT Region 1..365 /label= M4
XX FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
XX FT Disulfide-bond 4..85
XX FT Disulfide-bond 25..67
XX FT Disulfide-bond 56..80
XX FT Disulfide-bond 102..233
XX FT Disulfide-bond 143..159
XX FT Disulfide-bond 151..222
XX FT Disulfide-bond 247..316
XX FT Disulfide-bond 279..295
XX FT Disulfide-bond 306..334
XX FT Region 366..371 /label= X1
XX FT Region 372..392 /label= Y1
XX
XX DE4323754-Cl.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Wnendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence .392 AA;
XX
Query Match 100.0%; Score 510; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHGYRGKASTDTWGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHGYRGKASTDTWGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
```

```
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
RESULT 23
AAR66244
ID AAR66244 standard; protein; 393 AA.
XX AC AAR66244;
XX
XX 25-MAR-2003 (updated)
XX DT 22-AUG-1995 (first entry)
XX DE Bifunctional urokinase variant M11.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX KW urokinase; variant; mutein.
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX FT Region 1..365 /label= M4
XX FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
XX FT Disulfide-bond 4..85
XX FT Disulfide-bond 25..67
XX FT Disulfide-bond 56..80
XX FT Disulfide-bond 102..233
XX FT Disulfide-bond 143..159
XX FT Disulfide-bond 151..222
XX FT Disulfide-bond 247..316
XX FT Disulfide-bond 279..295
XX FT Disulfide-bond 306..334
XX FT Region 366..374 /label= X1
XX FT Region 375..393 /label= Y1
XX
XX DE4323754-Cl.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Wnendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 393 AA;
XX
Query Match 100.0%; Score 510; DB 16; Length 393;
```

Best Local Similarity 100.0%; Pred. No. 4.7e-41; Indels 0; Gaps 0;  
Matches 88; Conservative 0; Mismatches 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLIQOQTYHAHRSDALQGLGKHNYCRNPDN 60  
Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLIQOQTYHAHRSDALQGLGKHNYCRNPDN 61  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 24  
AAR66249  
ID AAR66249 standard; protein; 393 AA.  
XX AC AAR66249;  
XX XX  
DT 25-MAR-2003 (updated)  
DT 17-AUG-1995 (first entry)  
XX DE Bifunctional urokinase variant M16.  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
XX urokinase; variant; mutein.  
XX OS Synthetic.  
XX XX  
FH Key Location/Qualifiers  
FT Region 1..365 /label= M4  
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
FT Disulfide-bond 4..85  
FT Disulfide-bond 25..67  
FT Disulfide-bond 56..80  
FT Disulfide-bond 102..233  
FT Disulfide-bond 143..159  
FT Disulfide-bond 151..222  
FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Region 366..372 /label= X1  
FT Region 373..393 /label= Y1  
XX DE4323754-Cl.  
XX XX  
XX 01-DEC-1994. 93DE-4323754.  
XX 15-JUL-1993; 93DE-4323754.  
XX 15-JUL-1993; 93DE-4323754.  
XX (CHEF ) GRUENENTHAL GMBH.  
XX XX  
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;  
XX Whendt S;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with  
XX improved fibrinolytic and thrombin inhibiting activities, for  
XX treating cardiac and cerebral infarct, pulmonary embolism, etc  
XX  
XX Example 1; Page 10 and Fig 1; 34pp; German.  
XX PS  
XX Bifunctional urokinase derivatives corresponding to the formula  
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266  
XX are specific examples of such derivs. which have both improved  
XX fibrinolytic and thrombin-inhibiting activities, compared to known  
XX plasminogen activators or thrombin inhibitors. The proteins are  
XX useful as thrombolytic agents, e.g. for treatment of arterial

occlusions, deep vein thrombosis, cardiac and cerebral infarction  
and pulmonary embolism.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX Sequence 393 AA;  
Query Match 100.0%; Score 510; DB 16; Length 393;  
Best Local Similarity 100.0%; Pred. No. 4.7e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLIQOQTYHAHRSDALQGLGKHNYCRNPDN 60  
Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLIQOQTYHAHRSDALQGLGKHNYCRNPDN 61  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 25  
AAR66250  
ID AAR66250 standard; protein; 393 AA.  
XX AC AAR66250;  
XX XX  
DT 25-MAR-2003 (updated)  
DT 17-AUG-1995 (first entry)  
XX DE Bifunctional urokinase variant M17.  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
XX urokinase; variant; mutein.  
XX OS Synthetic.  
XX XX  
FH Key Location/Qualifiers  
FT Region 1..365 /label= M4  
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
FT Disulfide-bond 4..85  
FT Disulfide-bond 25..67  
FT Disulfide-bond 56..80  
FT Disulfide-bond 102..233  
FT Disulfide-bond 143..159  
FT Disulfide-bond 151..222  
FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Region 366..372 /label= X1  
FT Region 373..393 /label= Y1  
XX DE4323754-Cl.  
XX XX  
XX 01-DEC-1994. 93DE-4323754.  
XX 15-JUL-1993; 93DE-4323754.  
XX 15-JUL-1993; 93DE-4323754.  
XX (CHEF ) GRUENENTHAL GMBH.  
XX XX  
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;  
XX Whendt S;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with  
XX improved fibrinolytic and thrombin inhibiting activities, for  
XX treating cardiac and cerebral infarct, pulmonary embolism, etc  
XX  
XX Example 1; Page 10 and Fig 1; 34pp; German.  
XX PS

XX Bifunctional urokinase derivatives corresponding to the formula  
CC M4-XI-Y1 are claimed (see features table). Sequences AAR66244-R66266  
CC are specific examples of such derivs. which have both improved  
CC fibrinolytic and thrombin-inhibiting activities, compared to known  
CC plasminogen activators or thrombin inhibitors. The proteins are  
CC useful as thrombolytic agents, e.g. for treatment of arterial  
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction  
CC and pulmonary embolism.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 393 AA;  
  
Query Match 100.0%; Score 510; DB 16; Length 393;  
Best Local Similarity 100.0%; Pred. NO. 4.7e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 60  
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 61  
  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
  
RESULT 26  
AAR66251  
ID AAR66251 standard; protein; 393 AA.  
XX  
AC AAR66251;  
XX  
DT 25-MAR-2003 (updated)  
DT 17-AUG-1995 (first entry)  
XX  
DE Bifunctional urokinase variant M18.  
XX  
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
KW urokinase; variant; mutein.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Region 1..365  
FT /label= M4  
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
FT Disulfide-bond 4..85  
FT Disulfide-bond 25..67  
FT Disulfide-bond 56..80  
FT Disulfide-bond 102..233  
FT Disulfide-bond 143..159  
FT Disulfide-bond 151..222  
FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Disulfide-bond 366..372  
FT Region /label= X1  
FT Region /label= Y1  
XX  
PN DE4323754-C1.  
XX  
PD 01-DEC-1994.  
XX  
PP 15-JUL-1993; 93DE-4323754.  
XX  
PR 15-JUL-1993; 93DE-4323754.  
XX  
PA (CHEF ) GRUENENTHAL GMBH.  
XX  
PI Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;  
PI Wnendt S;  
XX

DR WPI; 1995-015191/03.  
XX  
PT New bifunctional urokinase derive and related plasmids - with  
PT improved fibrinolytic and thrombin inhibiting activities, for  
PT treating cardiac and cerebral infarct, pulmonary embolism, etc  
XX  
PS Example 1; Page 10 and Fig 1; 34pp; German.  
XX  
CC Bifunctional urokinase derivatives corresponding to the formula  
CC M4-XI-Y1 are claimed (see features table). Sequences AAR66244-R66266  
CC are specific examples of such derivs. which have both improved  
CC fibrinolytic and thrombin-inhibiting activities, compared to known  
CC plasminogen activators or thrombin inhibitors. The proteins are  
CC useful as thrombolytic agents, e.g. for treatment of arterial  
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction  
CC and pulmonary embolism.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 393 AA;  
  
Query Match 100.0%; Score 510; DB 16; Length 393;  
Best Local Similarity 100.0%; Pred. NO. 4.7e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 60  
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 61  
  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
  
RESULT 27  
AAR66252  
ID AAR66252 standard; protein; 393 AA.  
XX  
AC AAR66252;  
XX  
DT 25-MAR-2003 (updated)  
DT 17-AUG-1995 (first entry)  
XX  
DE Bifunctional urokinase variant M19.  
XX  
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
KW urokinase; variant; mutein.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Region 1..365  
FT /label= M4  
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
FT Disulfide-bond 4..85  
FT Disulfide-bond 25..67  
FT Disulfide-bond 56..80  
FT Disulfide-bond 102..233  
FT Disulfide-bond 143..159  
FT Disulfide-bond 151..222  
FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Disulfide-bond 366..372  
FT Region /label= X1  
FT Region /label= Y1  
XX  
PN DE4323754-C1.  
XX  
PD 01-DEC-1994.  
XX  
PP 15-JUL-1993; 93DE-4323754.  
XX

```

PR 15-JUL-1993; 93DE-4323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Wnendt S;
XX
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 393 AA;
XX
XX Query Match 100.0%; Score 510; DB 16; Length 393;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-41;
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
XX 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
XX
XX 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
XX 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
XX
XX RESULT 28
XX AAR66253
XX ID AAR66253 standard; protein; 393 AA.
XX
XX AC AAR66253;
XX
XX 25-MAR-2003 (updated)
XX DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M20.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Region /label= M4
XX /note= "unglycosylated prourokinase (Ser47-Leu411)"
XX
XX Disulfide-bond 1..365
XX Disulfide-bond 25..67
XX Disulfide-bond 56..80
XX Disulfide-bond 102..233
XX Disulfide-bond 143..159
XX Disulfide-bond 151..222
XX Disulfide-bond 247..316
XX Disulfide-bond 279..295
XX Disulfide-bond 306..334
XX Region 366..372
XX /label= X1
XX 373..393
XX Region /label= Y1

```

```

XX DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Wnendt S;
XX
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 393 AA;
XX
XX Query Match 100.0%; Score 510; DB 16; Length 393;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-41;
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
XX 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
XX
XX 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
XX 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
XX
XX RESULT 29
XX AAR99885
XX ID AAR99885 standard; peptide; 393 AA.
XX
XX AC AAR99885;
XX
XX 27-JAN-1997 (first entry)
XX
XX M36: fibrinolytic and anticoagulant activity contg. protein.
XX
XX Thrombin; factor Xa; anticoagulant; thrombolytic; plasminogen;
XX urokinase; activator; streptokinase; staphylokinase; APSAC;
XX anisolated plasminogen streptokinase activator complex; hirudin;
XX hirullin; antistatin; pWLT27; pWS1; pSE8; pW56.
XX
XX Synthetic.
XX
XX EP712934-A2.
XX
XX 22-MAY-1996.
XX
XX 03-NOV-1995; 95EP-0117316.
XX
XX 17-NOV-1994; 94DE-4440892.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX

```



```

XX Heinzel-Wieland R, Steffens GJ, Wnendt S;
XX WPI; 1996-240720/25.
XX
XX Proteins with fibrinolytic and anticoagulant activity - useful as
XX thrombolytic agents
XX
XX Disclosure; Fig 18; 59pp; German.
XX
XX New peptides (I) with fibrinolytic and anticoagulant activity
XX comprise a plasminogen-activating amino acid sequence (A) fused
XX at the N- and/or C-terminus to a thrombin and/or factor Xa
XX inhibiting amino acid sequence (B). Excluded from the claims
XX are (I) where (A) is Ser47 to Leu411 of unglycosylated urokinase
XX linked at the C-terminus to sequences (i) to (iii):
XX   TI-RP-T2-GGGGNGDFEIPBEYL-T3      (i)
XX   TI-RPFLRNPNKDYEPFWEDEKNE        (ii)
XX   TI-RPSSSEPEFEIDEEK               (iii)
XX   Where T1= P or V; T2 = L or a bond; T3= Q or OH.
XX   (A) is pref. (pro)urokinase; tissue plasminogen activator (tPA),
XX   bat-PA (all opt. modified by deletion, substitution, insertion and/or
XX   addn.); streptokinase; staphylokinase; and/or APSAC (anisolated
XX   plasminogen streptokinase activator complex), esp. prourokinase
XX   (411 amino acids) or its Ser47 to Leu411 or Ser138 to Leu411
XX   fragments, or t-PA (527 amino acids) or its Ser89Arg to 527Pro or
XX   174Ser to 527Pro fragments.
XX   (B) has hirudin or hirullin activities; or is derived from the human
XX   thrombin receptor, antistasin and/or the tick anticoagulant peptide.
XX   Most pref. are the 65 amino acid hirudin sequence or one of the six
XX   sequences given in AAR99879 to AAR99884.
XX   Plasmids pWLT27 (M51), pWS1 (M5112), pSE8 (M36) and pW56 (M43)
XX   contain the sequences encoding AAR99885 to AAR99889, respectively.
XX   The products were tested in human citrated plasma (5 microg in 200
XX   microl 1:10 diluted plasma). The thrombin time was then 1.2, 3, 2.8
XX   and 1.2 times greater, respectively, than in the absence of the product.
XX
XX   Sequence 393 AA;
SQ
  Query Match          100.0%; Score 510; DB 17; Length 393;
  Best Local Similarity 100.0%; Pred. No. 4.7e-41;
  Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 60
Db 3 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 62

QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
Db 63 RRRPWCYVQVGLKPLVQECWVHDCADGK 90

RESULT 30
AAR99596
ID AAR99596 standard; Protein; 393 AA.
AC
XX AAR99596;
XX
XX 05-DEC-1996 (first entry)
XX
XX Chimeric protein M37 encoded by pSE9.
XX
XX Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;
XX plasminogen activating sequence; fibrinolysis; infarction;
XX angina pectoris; deep vein thrombosis.
XX
XX Synthetic.
XX
XX EP714982-A2.
XX
XX 05-JUN-1996.
XX
XX 16-NOV-1995; 95EP-0118050.

```

```

XX 30-NOV-1994; 94DE-4442665.
XX (CHEF ) GRUENENTHAL GMEH.
XX
XX Heinzel-Wieland R, Janocha E, Steffens GJ, Wnendt S;
XX WPI; 1996-269715/28.
XX
XX Chimeric protein contg. plasminogen activating sequence and
XX thrombin-inhibiting sequence - useful as thrombus-specific
XX thrombolytic agent with rapid action
XX
XX Example 1; Page 19-20; 37pp; German.
XX
XX Example 1 describes the prodn. of plasmids pSE1 and pSE9
XX contg. a DNA encoding a chimeric protein with fibrinolytic and
XX thrombin-inhibiting properties.
XX pSE1 encodes the protein given in AAR99597 and pSE9 encodes
XX the protein given in AAR99596.
XX
XX Sequence 393 AA;
SQ
  Query Match          100.0%; Score 510; DB 17; Length 393;
  Best Local Similarity 100.0%; Pred. No. 4.7e-41;
  Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 60
Db 3 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRPN 62

QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
Db 63 RRRPWCYVQVGLKPLVQECWVHDCADGK 90

Search completed; December 3, 2003, 14:39:09
Job time : 14.8621 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:23 ; Search time 10.8276 Seconds  
(without alignments)  
2097.294 Million cell updates/sec

Title: US-09-880-503-1

Perfect score: 510

Sequence: 1 KTCYEGNGHFVRGKASTDTM.....QVGLKPLVQECMWHDCADGK 88

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	495	97.1	154	4 Q96SE8	Q96SE8 homo sapien
2	415	81.4	433	6 Q8MIL0	Q8MIL0 oryctolagus
3	415	81.4	433	6 Q8MHY7	Q8MHY7 oryctolagus
4	396	77.6	157	6 Q9TVA8	Q9TVA8 bos taurus
5	387	75.9	231	11 Q8C6L2	Q8C6L2 mus musculus
6	313	61.4	214	6 Q9XT70	Q9XT70 oryctolagus
7	235	46.1	564	6 Q8KBE1	Q8KBE1 oryctolagus
8	226	44.3	516	4 Q9BU99	Q9BU99 homo sapien
9	219	42.9	559	11 Q91VP2	Q91VP2 mus musculus
10	218.5	42.8	90	4 Q8NG20	Q8NG20 homo sapien
11	215	42.2	395	4 Q9BZW1	Q9BZW1 homo sapien
12	210	41.2	562	6 Q8SQ23	Q8SQ23 mus musculus
13	209.5	41.1	517	11 Q8KOD2	Q8KOD2 mus musculus
14	193.5	37.9	560	4 Q14520	Q14520 homo sapien
15	191	37.5	202	13 Q90675	Q90675 gallus gall
16	188	36.9	653	11 Q8VCS4	Q8VCS4 mus musculus

17	186.5	36.6	597	11 Q35727	Q35727 mus musculus
18	177.5	34.8	616	6 Q97507	Q97507 sus scrofa
19	172.5	33.8	421	13 Q8AXX3	Q8AXX3 xenopus lae
20	170.5	33.4	615	4 Q81ZZ5	Q81ZZ5 homo sapien
21	163.5	32.1	947	13 Q8AXY6	Q8AXY6 gallus gall
22	158	31.0	452	13 Q90Y90	Q90Y90 xenopus lae
23	158	31.0	454	6 Q46506	Q46506 papio hamad
24	156	30.6	113	4 Q9UIR5	Q9UIR5 homo sapien
25	154	30.2	806	6 Q18783	Q18783 macropus eu
26	153	30.0	113	4 Q9UIR7	Q9UIR7 homo sapien
27	151.5	29.7	359	6 Q8WMR1	Q8WMR1 canis famil
28	149.5	29.3	812	11 Q9R0W3	Q9R0W3 rattus norv
29	149	29.2	716	13 Q91691	Q91691 xenopus lae
30	148.5	29.1	381	4 Q8N2U4	Q8N2U4 homo sapien
31	148.5	29.1	399	4 Q96GL8	Q96GL8 homo sapien
32	148.5	29.1	420	4 Q9BTP9	Q9BTP9 homo sapien
33	148.5	29.1	424	4 Q8NCW1	Q8NCW1 homo sapien
34	148.5	29.1	462	4 Q8NCW0	Q8NCW0 homo sapien
35	148.5	29.1	801	11 Q8KQ08	Q8KQ08 mus musculus
36	148.5	29.1	944	11 Q8C3W2	Q8C3W2 mus musculus
37	148.5	29.1	944	11 Q8BSP6	Q8BSP6 mus musculus
38	148	29.0	105	4 Q9UIR8	Q9UIR8 homo sapien
39	147	28.8	113	4 Q9UIR6	Q9UIR6 homo sapien
40	147	28.8	648	4 Q9H1V4	Q9H1V4 homo sapien
41	146.5	28.7	461	11 Q8K1S7	Q8K1S7 mus musculus
42	146	28.6	393	4 Q9BRB6	Q9BRB6 homo sapien
43	146	28.6	567	4 Q13208	Q13208 homo sapien
44	146	28.6	937	11 Q8BNP9	Q8BNP9 mus musculus
45	146	28.6	937	11 Q8BG10	Q8BG10 mus musculus

#### ALIGNMENTS

RESULT 1

Q96SE8 ID Q96SE8 PRELIMINARY; PRT; 154 AA.  
AC Q96SE8;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Urokinase-type plasminogen activator amino-terminal fragment.

GN ATF.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Fu J., Bai X., Ruan C.;

RT "Cloning and expression of the amino-terminal fragment of human

urokinase-type plasminogen activator.";

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Bai X., Fu J., Wang W., Xi X., Ruan C.;

RT "Overexpression of the amino-terminal fragment of human urokinase-type

plasminogen activator in breast cancer cells results in decreased

tumor invasion, growth and angiogenesis.";

Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

-1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

EMBL; AY029537; AAK38734.1; -.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR000001; Kringle.

DR Pfam; PF00051; kringle\_1.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00181; EGF; 1.

DR SMART; SM00130; KR; 1.

DR PROSITE; PS00022; EGF\_1; 1.

DR PROSITE; PS00021; KRINGLE\_1; 1.

DR PROSITE; PS00070; KRINGLE\_2; 1.

```
KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CCF2FCFF505572 CRC64;

Query Match      97.1%; Score 495; DB 4; Length 154;
Best Local Similarity 98.9%; Pred. No. 1.6e-51;
Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHN YCRNPDN 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
68 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHN YCRNPDN 127
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADG 154
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 2
Q8MHY7
ID Q8MHY7 PRELIMINARY; PRT; 433 AA.
AC Q8MHY7;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RL "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY029517; AAK40239.1; -.
DR EMBL; AB087224; BAC02685.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Serprotease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match      81.4%; Score 415; DB 6; Length 433;
Best Local Similarity 79.5%; Pred. No. 2e-41;
Matches 70; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHN YCRNPDN 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
70 KTCYHGDSYRGKANTDMDRPLAWN SANVLTXTYHAHRPDLQLGLGKHN YCRNPDH 129
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 88
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 130 QRRPWCYVQVGLKQLIQECKVHDCSSGK 157
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
Q9TVAS
ID Q9TVAS PRELIMINARY; PRT; 157 AA.
AC Q9TVAS;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
```

```

OX NCBI_TaxID=9913;
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21071388; PubMed=11204721;
RA Balcerzak D., Querengesser L., Dixon W.T., Baracos V.B.;
RT "Coordinate expression of matrix-degrading proteinases and their
RT activators and inhibitors in bovine skeletal muscle.";
RL J. Anim. Sci. 79:94-107(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF144761; AAD30301.1; -.
DR HSSP; P00749; IURK.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; P00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kinase; Kringle.
FT NON_TER 1
FT NON_TER 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 77.6%; Score 396; DB 6; Length 157;
Best Local Similarity 76.1%; Pred. No. 1.2e-39;
Matches 67; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFRGKASTDTMGKPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db
36 KTCYQNGHSYRGKANDLSGRPCLDWSPVLLKWKYHAHRSDALQLGLGKHNYCRNPDN 95
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db
96 QRRPWCYVQIGLQKQFVQFCMVQDCSVGK 123

RESULT 5
Q8C6L2 PRELIMINARY; PRT; 231 AA.
AC Q8C6L2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK054349; BACJ5743.1; -.
FT NON_TER 231
FT NON_TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25E8980A682737F2 CRC64;

Query Match 75.9%; Score 387; DB 11; Length 231;
Best Local Similarity 72.7%; Pred. No. 2.3e-38;
Matches 64; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFRGKASTDTMGKPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db
69 KTCYHGNGDSYRGKANDTKGRPCLDWSPVLLKWKYHAHRSDALQLGLGKHNYCRNPDN 128
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88

Db
129 QRRPWCYVQIGLQKQFVQFCMVHDCSLSK 156

RESULT 6
Q9XT70 PRELIMINARY; PRT; 214 AA.
AC Q9XT70;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yin J., Idell S.;
RT "partial mRNA of rabbit uPA.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; AAD39351.1; -.
DR HSSP; P00749; IELN.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Serprotease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
FT NON_TER 214
FT NON_TER 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 61.4%; Score 313; DB 6; Length 214;
Best Local Similarity 78.3%; Pred. No. 1.6e-29;
Matches 54; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 20 MGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVQVGLKPLVQEC 79
Db
1 MDRPCLAWSANVLTKTYHAHRPDALQLGLGKHNYCRNPDHQRPRPWCYVQVGLKPLVQEC 60
QY 80 MVHDCADGK 88
Db
61 KVHDSGKK 69

RESULT 7
Q8MKB1 PRELIMINARY; PRT; 564 AA.
AC Q8MKB1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tissue-type plasminogen activator.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
```

```
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AY029518; AAK40240.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEFG.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 564 AA; 62726 MW; 459D8BAC6D4A937C CRC64;

Query Match 46.1%; Score 235; DB 6; Length 564;
Best Local Similarity 50.6%; Pred. No. 1e-19;
Matches 42; Conservative 10; Mismatches 31; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDR 62
Db 217 CYLKGQAYRGTHSLTTSASCLPWN SMLLVGEKYTARQSNAELGLGKHNYCRNP 276

QY 63 RFWCYVQVGLKPLVQECMVHDC 85
Db 277 KPWCHVKNRKLVI EYCDVPQCA 299

RESULT 8
Q9BU99 PRELIMINARY; PRT; 516 AA.
AC Q9BU99;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DDAJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC002795; AAH02795.1; -.
DR HSP; P00750; IASH.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEFG.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00051; kringle; 2.
```

```
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match 44.3%; Score 226; DB 4; Length 516;
Best Local Similarity 47.7%; Pred. No. 1.1e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDR 61
Db 80 TCYEDQGISYRGTWSTASGAECTWNNSALAQKPYSGRRPD ARLGLGNHNYCRNPDR 139

QY 62 RFPWCYVQVGLKPLVQECMVHDC 87
Db 140 SKPWCYVFKAGYSSEFCSTPACSE 165

RESULT 9
Q91VP2 PRELIMINARY; PRT; 559 AA.
AC Q91VP2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
GN PLAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDAJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC011256; AAH11256.1; -.
DR HSP; P00761; 1ANI.
DR MGD; MGI:97610; Plat.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEFG.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR PROSITE; PS00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
```



```
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00039; fn1; 1.
DR Pfam: PF00051; kringle; 2.
DR Pfam: PF00089; trypsin; 1.
DR ProDom: PD000395; Kringle; 2.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00058; fn1; 1.
DR SMART: SM00130; KR; 2.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 2.
DR PROSITE: PS00070; KRINGLE_2; 2.
DR PROSITE: PS02040; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase;
KW Serine protease.
SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;

Query Match 41.2%; Score 210; DB 6; Length 562;
Best Local Similarity 43.7%; Pred. No. 1e-16;
Matches 38; Conservative 9; Mismatches 40; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNR 61
Db 126 TCYEGNGHYRGKASDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNR 61
QY 62 RRPWCYVQVGLKPLVQECMVHDCADCK 88
Db 186 SRPWCYIFKAEKXSPDFCSTPACTKEK 212

RESULT 13
Q8K0D2 PRELIMINARY; PRT; 517 AA.
AC Q8K0D2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC031775; AAH31775.1; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR006209; EGF-like.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 3.
```

```
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match 41.1%; Score 209.5; DB 11; Length 517;
Best Local Similarity 45.8%; Pred. No. 1.1e-16;
Matches 38; Conservative 13; Mismatches 31; Indels 1; Gaps 1;

QY 3 CYEGNGHYRGKASDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNR 62
Db 151 CYVGDGYSYRGKSVKTVNQNPCLYWN SHLLQETYNFMEDAETHGIAEHFRCRNPDDGH 210
QY 63 RPPWCYVQVGLKPLVQECMVHDC 84
Db 211 KPCFCYKVNSEKVKMEYCDVTVC 233

RESULT 14
Q14520 PRELIMINARY; PRT; 560 AA.
AC Q14520; O00663;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HGF activator like protein (Hyaluronan binding protein 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RA Kitamura N.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2];
RP SEQUENCE FROM N.A.
RX MEDLINE=96425001; PubMed=8827452;
RA Choi-Mura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
RA Tomita M.;
RT "Purification and characterization of a novel hyaluronan-binding
RT protein (PHBP) from human plasma: it has three EGF, a kringle and a
RT serine protease domain, similar to hepatocyte growth factor
RT activator.";
RL J. Biochem. 119:1157-1165(1996).
RN [3];
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; D49742; BAA08576.1; -.
DR EMBL; S83182; AAB46909.1; -.
DR EMBL; BC031412; AAH31412.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.033; -.
DR Genew; HGNC:4798; HABP2.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR006209; EGF like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 3.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF_1; 3.
```

Qy	63	RPWCYVQVGLKPLVQSCMWHDC	84
	:		:
Db	101	QPCHWWKORQLTWYCDVPQC	122
	:		:

  

RESULT 16	
Q8VCS4	
ID	Q8VCS4 PRELIMINARY; PRT; 653 AA.
Q8VCS4:	
AC	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Hypothetical 70.6 kDa protein.
DE	Mus musculus (Mouse).
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX	NCBI_TaxID=10090;
OX	[1]
RP	SEQUENCE FROM N.A.
RP	RC TISSUE=Liver;
RC	Strasbourg R.;
RA	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC	-1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC	EMBL; BC019376; AAH19376.1; -.
DR	HSP; P00761; IAN1.
DR	InterPro; IPR001314; Chymotrypsin.
DR	InterPro; IPR000742; EGF 2.
DR	InterPro; IPR006209; EGF like.
DR	InterPro; IPR000083; Fibrinctn1.
DR	InterPro; IPR000562; FN Type_II.
DR	InterPro; IPR006210; EGF.
DR	InterPro; IPR000001; Kringle.
DR	InterPro; IPR001254; Ser_protease_Try.
DR	Pfam; PF00008; EGF; 2.
DR	Pfam; PF00039; fn1; 1.
DR	Pfam; PF00040; fn2; 1.
DR	Pfam; PF00051; kringle; 1.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	PRINTS; PR00013; FNTYPEII.
DR	PRINTS; PR00018; KRINGLE.
DR	ProDom; PD000995; FN Type_II; 1.
DR	ProDom; PD000395; Kringle; 1.
DR	SMART; SM00181; EGF; 2.
DR	SMART; SM00059; FN2; 1.
DR	SMART; SM00130; KR; 1.
DR	SMART; SM00020; Tryp_Spc; 1.
DR	PROSITE; PS00022; EGF 1; 2.
DR	PROSITE; PS01186; EGF 2; 1.
DR	PROSITE; PS01253; FIBRONECTIN 1; 1.
DR	PROSITE; PS00023; FIBRONECTIN 2; 1.
DR	PROSITE; PS00021; KRINGLE 1; 1.
DR	PROSITE; PS00070; KRINGLE 2; 1.
DR	PROSITE; PS00240; TRYPSIN DOM; 1.
DR	PROSITE; PS00134; TRYPSIN HIS; 1.
DR	PROSITE; PS00135; TRYPSIN SER; 1.
KW	Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW	Kringle; Protease; Serine protease.
SQ	SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

  

Query Match	36.9%	Score 188;	DB 11;	Length 653;
Best Local Similarity	53.0%;	Pred.No. 5.3e-14;		
Matches 35;	Conservative 6;	Mismatches 25;	Indels 0;	Gaps 0;

  

Qy	3	CVEGNHGYRCASTDTMGRCPLPWNSATVLOQTVAHRSDALQLGLGKHVYCRNPDRR	62
	:		:
Db	293	CFLGNNGTYRGVASTAAGSLGSLANNSDLLYQELHVDVSAAAVLLGLGFHAYCRNPDKDE	342
	:		:

  

Qy	63	RPWCYV	68
	:		:
Db	343	RPWCYV	348
	:		:



RESULT 17	RESULT 18
<p> O35727  ID O35727 PRELIMINARY; PRT; 597 AA.  AC O35727  DT 01-JAN-1998 (T-EMBLrel. 05, Created)  DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)  DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  DE Factor XII.  GN F12.  OS Mus musculus (Mouse).  OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  OX NCBI_TaxID=10090;  RN [1]  RP SEQUENCE FROM N.A.  RC TISSUE=Liver;  RT "Porcine liver factor XII.";  RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  DR EMBL; X99571; CAA67891.1; -.  DR HSP; P00760; IAQ7.  DR MEROPS; S01.211; -.  DR MG; MGI:1891012; F12.  DR InterPro; IPR001314; Chymotrypsin.  DR InterPro; IPR006209; EGF like.  DR InterPro; IPR000083; Fibnctnl.  DR InterPro; IPR000562; FN Type_II.  DR InterPro; IPR006210; IEGF.  DR InterPro; IPR000001; Kringle.  DR InterPro; IPR001254; Ser_protease_Try.  DR Pfam; PF00008; EGF; 2.  DR Pfam; PF00039; fn1; 1.  DR Pfam; PF00040; fn2; 1.  DR Pfam; PF00051; kringle; 1.  DR Pfam; PF00089; trypsin; 1.  DR PRINTS; PR00722; CHYMOTRYPSPIN.  DR PRINTS; PR00013; FNTYPEII.  DR ProDom; PD000995; FN Type II; 1.  DR ProDom; PD000395; Kringle; 1.  DR SMART; SM00181; EGF; 2.  DR SMART; SM00058; FN1; 1.  DR SMART; SM00059; FN2; 1.  DR SMART; SM00130; KR; 1.  DR SMART; SM00020; Tryp_Spc; 1.  DR PROSITE; PS00022; EGF_1; 2.  DR PROSITE; PS01186; EGF_2; 1.  DR PROSITE; PS01253; FIBRONECTIN_1; 1.  DR PROSITE; PS00023; FIBRONECTIN_2; 1.  DR PROSITE; PS00021; KRINGLE_1; 1.  DR PROSITE; PS00070; KRINGLE_2; 1.  DR PROSITE; PS00240; TRYPSPIN_DOM; 1.  DR PROSITE; PS00134; TRYPSPIN_HIS; 1.  DR PROSITE; PS00135; TRYPSPIN_SER; 1.  KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;  KW Serine protease.  SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;  </p> <p> Query Match 36.6%; Score 186.5; DB 11; Length 597;  Best Local Similarity 44.0%; Pred. No. 7.2e-14;  Matches 37; Conservative 8; Mismatches 34; Indels 5; Gaps 2;  </p> <p> QY 2 TCYEGNGHFRGKASDTMGRCPLPNSATVLQOTY-HAHRSDALQLGKHNCRNPDN 60  DB 216 TCYEGRLSVRGAGTGTQSGACQVR-----TVEATYRNWTEKQSLWGIGHAFRCRNP 271  </p> <p> QY 61 RRPWCYVQVGLKPLVQECWVHDC 84  DB 272 DTRPWCFCVSGDELSDWYCGLEQC 295  </p>	<p> O97507  ID O97507 PRELIMINARY; PRT; 616 AA.  AC O97507  DT 01-MAY-1999 (T-EMBLrel. 10, Created)  DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)  DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  DE FXII.  OS Sus scrofa (Pig).  OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  OX NCBI_TaxID=9823;  RN [1]  RP SEQUENCE FROM N.A.  RC TISSUE=Liver;  RT Takahashi T.; Kihara T.;  RL "Porcine liver factor XII.";  RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  DR EMBL; AB022426; BAA37148.1; -.  DR HSP; P00763; IDPO.  DR MEROPS; S01.211; -.  DR InterPro; IPR001314; Chymotrypsin.  DR InterPro; IPR000742; EGF 2.  DR InterPro; IPR006209; EGF like.  DR InterPro; IPR000083; Fibnctnl.  DR InterPro; IPR000562; FN Type_II.  DR InterPro; IPR006210; IEGF.  DR InterPro; IPR000001; Kringle.  DR InterPro; IPR001254; Ser_protease_Try.  DR Pfam; PF00008; EGF; 2.  DR Pfam; PF00039; fn1; 1.  DR Pfam; PF00040; fn2; 1.  DR Pfam; PF00051; kringle; 1.  DR Pfam; PF00089; trypsin; 1.  DR PRINTS; PR00722; CHYMOTRYPSPIN.  DR PRINTS; PR00013; FNTYPEII.  DR ProDom; PD000995; FN Type II; 1.  DR ProDom; PD000395; Kringle; 1.  DR SMART; SM00181; EGF; 2.  DR SMART; SM00058; FN1; 1.  DR SMART; SM00059; FN2; 1.  DR SMART; SM00130; KR; 1.  DR SMART; SM00020; Tryp_Spc; 1.  DR PROSITE; PS00022; EGF_1; 2.  DR PROSITE; PS01186; EGF_2; 1.  DR PROSITE; PS01253; FIBRONECTIN_1; 1.  DR PROSITE; PS00023; FIBRONECTIN_2; 1.  DR PROSITE; PS00021; KRINGLE_1; 1.  DR PROSITE; PS00070; KRINGLE_2; 1.  DR PROSITE; PS00240; TRYPSPIN_DOM; 1.  DR PROSITE; PS00134; TRYPSPIN_HIS; 1.  DR PROSITE; PS00135; TRYPSPIN_SER; 1.  KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;  KW Serine protease.  SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBD1A9 CRC64;  </p> <p> Query Match 34.8%; Score 177.5; DB 6; Length 616;  Best Local Similarity 44.6%; Pred. No. 9e-13;  Matches 37; Conservative 6; Mismatches 35; Indels 5; Gaps 2;  </p> <p> QY 3 CYEGNGHFRGKASDTMGRCPLPNSATVLQOTY-HAHRSDALQLGKHNCRNPDN 61  DB 217 CYSDELGLSVRGAGTGTLSGAPQWAS----EATYWNNTAEQALNWGLGIGHAFRCRNP 272  </p> <p> QY 62 RRPWCYVQVGLKPLVQECWVHDC 84  DB 273 TRPWCFCVSGDELSDWYCGLEQC 295  </p>

RESULT 17	RESULT 18
<p> O35727  ID O35727 PRELIMINARY; PRT; 597 AA.  AC O35727  DT 01-JAN-1998 (T-EMBLrel. 05, Created)  DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)  DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  DE Factor XII.  GN F12.  OS Mus musculus (Mouse).  OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  OX NCBI_TaxID=10090;  RN [1]  RP SEQUENCE FROM N.A.  RC TISSUE=Liver;  RT "Porcine liver factor XII.";  RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  DR EMBL; X99571; CAA67891.1; -.  DR HSP; P00760; IAQ7.  DR MEROPS; S01.211; -.  DR MG; MGI:1891012; F12.  DR InterPro; IPR001314; Chymotrypsin.  DR InterPro; IPR006209; EGF like.  DR InterPro; IPR000083; Fibnctnl.  DR InterPro; IPR000562; FN Type_II.  DR InterPro; IPR006210; IEGF.  DR InterPro; IPR000001; Kringle.  DR InterPro; IPR001254; Ser_protease_Try.  DR Pfam; PF00008; EGF; 2.  DR Pfam; PF00039; fn1; 1.  DR Pfam; PF00040; fn2; 1.  DR Pfam; PF00051; kringle; 1.  DR Pfam; PF00089; trypsin; 1.  DR PRINTS; PR00722; CHYMOTRYPSPIN.  DR PRINTS; PR00013; FNTYPEII.  DR ProDom; PD000995; FN Type II; 1.  DR ProDom; PD000395; Kringle; 1.  DR SMART; SM00181; EGF; 2.  DR SMART; SM00058; FN1; 1.  DR SMART; SM00059; FN2; 1.  DR SMART; SM00130; KR; 1.  DR SMART; SM00020; Tryp_Spc; 1.  DR PROSITE; PS00022; EGF_1; 2.  DR PROSITE; PS01186; EGF_2; 1.  DR PROSITE; PS01253; FIBRONECTIN_1; 1.  DR PROSITE; PS00023; FIBRONECTIN_2; 1.  DR PROSITE; PS00021; KRINGLE_1; 1.  DR PROSITE; PS00070; KRINGLE_2; 1.  DR PROSITE; PS00240; TRYPSPIN_DOM; 1.  DR PROSITE; PS00134; TRYPSPIN_HIS; 1.  DR PROSITE; PS00135; TRYPSPIN_SER; 1.  KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;  KW Serine protease.  SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;  </p> <p> Query Match 36.6%; Score 186.5; DB 11; Length 597;  Best Local Similarity 44.0%; Pred. No. 7.2e-14;  Matches 37; Conservative 8; Mismatches 34; Indels 5; Gaps 2;  </p> <p> QY 2 TCYEGNGHFRGKASTDTMGRPCLPNSATVLOQTY-HAHRSDALQLGLGKHNYCRNPDN 60  DB 216 TCYEGRLSVRGAGTGTQSGACQVR-----TVEATYRNWTEKQSLWGSLGHAFRCRNP 271  </p> <p> QY 61 RRPWCYVQVGLKPLVQECWVHDC 84  DB 272 DTRPWCFCVSGDELSDWYCGLEQC 295  </p>	<p> O97507  ID O97507 PRELIMINARY; PRT; 616 AA.  AC O97507  DT 01-MAY-1999 (T-EMBLrel. 10, Created)  DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)  DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  DE FXII.  OS Sus scrofa (Pig).  OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  OX NCBI_TaxID=9823;  RN [1]  RP SEQUENCE FROM N.A.  RC TISSUE=Liver;  RT Takahashi T.; Kihara T.;  RL "Porcine liver factor XII.";  RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  DR EMBL; AB022426; BAA37148.1; -.  DR HSP; P00763; IDPO.  DR MEROPS; S01.211; -.  DR InterPro; IPR001314; Chymotrypsin.  DR InterPro; IPR000742; EGF 2.  DR InterPro; IPR006209; EGF like.  DR InterPro; IPR000083; Fibnctnl.  DR InterPro; IPR000562; FN Type_II.  DR InterPro; IPR006210; IEGF.  DR InterPro; IPR000001; Kringle.  DR InterPro; IPR001254; Ser_protease_Try.  DR Pfam; PF00008; EGF; 2.  DR Pfam; PF00039; fn1; 1.  DR Pfam; PF00040; fn2; 1.  DR Pfam; PF00051; kringle; 1.  DR Pfam; PF00089; trypsin; 1.  DR PRINTS; PR00722; CHYMOTRYPSPIN.  DR PRINTS; PR00013; FNTYPEII.  DR ProDom; PD000995; FN Type II; 1.  DR ProDom; PD000395; Kringle; 1.  DR SMART; SM00181; EGF; 2.  DR SMART; SM00058; FN1; 1.  DR SMART; SM00059; FN2; 1.  DR SMART; SM00130; KR; 1.  DR SMART; SM00020; Tryp_Spc; 1.  DR PROSITE; PS00022; EGF_1; 2.  DR PROSITE; PS01186; EGF_2; 1.  DR PROSITE; PS01253; FIBRONECTIN_1; 1.  DR PROSITE; PS00023; FIBRONECTIN_2; 1.  DR PROSITE; PS00021; KRINGLE_1; 1.  DR PROSITE; PS00070; KRINGLE_2; 1.  DR PROSITE; PS00240; TRYPSPIN_DOM; 1.  DR PROSITE; PS00134; TRYPSPIN_HIS; 1.  DR PROSITE; PS00135; TRYPSPIN_SER; 1.  KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;  KW Serine protease.  SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBD1A9 CRC64;  </p> <p> Query Match 34.8%; Score 177.5; DB 6; Length 616;  Best Local Similarity 44.6%; Pred. No. 9e-13;  Matches 37; Conservative 6; Mismatches 35; Indels 5; Gaps 2;  </p> <p> QY 3 CYEGNGHFRGKASTDTMGRPCLPNSATVLOQTY-HAHRSDALQLGLGKHNYCRNPDN 61  DB 217 CYSDRLSVRGAGTGTLSGAPQWAS----EATYWNNTAEQALNWGLGSHAFRCRNP 272  </p> <p> QY 62 RRPWCYVQVGLKPLVQECWVHDC 84  DB 273 TRPWCFCVSGDELSDWYCGLEQC 295  </p>

```

QBAXX3
ID QBAXX3 PRELIMINARY; PRT; 421 AA.
AC QBAXX3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Kremen2.
GN KRM2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22308873; PubMed=12421700;
RA Davidson G., Mao B., Del Barco Barrantes I., Niehrs C.;
RT "kremen proteins interact with Dickkopf1 to regulate anteroposterior
RL CNS patterning.";
RL Development 129:5587-5596(2002).
DR EMBL; AY150813; AAN64661.1; -.
SQ SEQUENCE 421 AA; 46274 MW; 74D04DB1682CD15F CRC64;

Query Match 33.8%; Score 172.5; DB 13; Length 421;
Best Local Similarity 50.7%; Pred. No. 2.4e-12;
Matches 34; Conservative 4; Mismatches 26; Indels 3; Gaps 2;

QY 3 CYEGNGHFYRGKAS-TDTMGRLPWN SATVLQQTYYAHRS DALQLGLGKHN YCRNPDR 61
DB 29 CTFVNGRDYRGTSVQAGPGTFLYWNQTT--QHL YNAQSDPDGELGLGNH YCRNPDA 86

QY 62 RRPWCYV 68
DB 87 VQPPWCYV 93

RESULT 20
QB1ZZ5
ID QB1ZZ5 PRELIMINARY; PRT; 615 AA.
AC QB1ZZ5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Coagulation factor XII-Mie.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wada H., Nishiocka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;
RT "Molecular characterization of coagulation factor XII-Mie.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095845; BAC23095.1; -.
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 33.4%; Score 170.5; DB 4; Length 615;
Best Local Similarity 48.5%; Pred. No. 6.2e-12;
Matches 33; Conservative 6; Mismatches 24; Indels 5; Gaps 2;

QY 2 TCYEGNGHFYRGKAS-TDTMGRLPWN SATVLQQTYYAHRS DALQLGLGKHN YCRNPDR 60
DB 216 SCYDGRGLSYRGUARTL TSGAPCQPWAS----EATYRNVTA EQARNWGLGGHAF CRNPDR 271

QY 61 RRPWCYV 68
DB 272 DIRPWCYV 279

RESULT 21
QBAXY6
ID QBAXY6 PRELIMINARY; PRT; 947 AA.
AC QBAXY6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mus musculus (Mus).
OS Mus musculus (Mus)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20538710; PubMed=11083926;
RA Ip F.C., Glass D.G., Gies D.R., Cheung J., Lai K.O., Fu A.X.;
RA Yancopoulos G.D., Ip N.Y.;
RT "Cloning and characterization of muscle-specific kinase in chicken.";
RL Mol. Cell. Neurosci. 16:661-673(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Gies D., Glass D.J., Yancopoulos G.D.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY143173; AAN05008.1; -.
KW Receptor; Kinase.
SQ SEQUENCE 947 AA; 105588 MW; E6C3FCC5796CC8BE CRC64;

Query Match 32.1%; Score 163.5; DB 13; Length 947;
Best Local Similarity 39.8%; Pred. No. 6.9e-11;
Matches 38; Conservative 6; Mismatches 27; Indels 27; Gaps 6;

QY 1 KTCYEGNGHFYRGKAS-TDTMGRLPWN SATVLQQTYYAHRS DALQLGLGK 51
DB 463 RTCYSGNGQYVQGWANVTASGIPCQKWS----DQAPHLHRRTPQVPPELSDA----- 510

QY 52 HNYCRNP-DNRRRPWCYVQGLKPLV--QECMVHDCAD 86
DB 511 ENYCRNFGENEPWCYTK---DPSVTWYCVSVSPCGD 545

RESULT 22
Q90Y90
ID Q90Y90 PRELIMINARY; PRT; 452 AA.
AC Q90Y90;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE KREMEN.
GN KREMEN.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21167372; PubMed=11267660;
RA Nakamura T., Aoki S., Kitajima K., Funakoshi H., Takahashi T.;
RA Matsumoto K., Nakamura T.;
RT "Molecular cloning and characterization of Kremen, a novel kring-
RL containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.
DR EMBL; AB070851; BAB64294.1; -.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR000001; Kringles.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringles; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringles; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.

```

```
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
SQ SEQUENCE 452 AA; 50188 MW; ED24BCD1AF4564E2 CRC64;

Query Match 31.0%; Score 158; DB 13; Length 452;
Best Local Similarity 44.1%; Pred. No. 1.4e-10;
Matches 30; Conservative 9; Mismatches 25; Indels 4; Gaps 2;

QY 3 CYEGNGHFGYRGKASTDTM--GRPCLPWNSATVLQQTYYHAHRSDALQGLGKHNYCRNPDN 60
DQ ||| ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
Db 30 CYTVNGADYRGTONQSLDGRKPLFWE--TFQHPYNTLKYENGSGGLGEHNYCRNPDG 87
QY 61 RRRPWCYV 68
DQ ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
Db 88 DVSPWCYI 95

RESULT 23
O46506 PRELIMINARY; PRT; 454 AA.
AC O46506;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein a (Fragment).
GN BABAPOA.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF029691; AAB97886.1; -.
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.999; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
KW Serine protease.
FT NON_TER 1
SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;

Query Match 31.0%; Score 158; DB 6; Length 454;
Best Local Similarity 39.8%; Pred. No. 1.4e-10;
Matches 35; Conservative 10; Mismatches 31; Indels 12; Gaps 4;

QY 3 CYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQ--QTYHAHRSDALQGLGKHNYCRNPDN 60
DQ ||| ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| ||| : : ||| |||
Db 102 CYHGGQSGYRGFSFTVTGRTCSQSSSWTFRHQKTPENHPNDGLTM-----NYCRNPD 156
QY 61 RRRPWCYVQGLKPLV--QECMVHDCAD 86
DQ ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
Db 61 RRRPWCYVQGLKPLV--QECMVHDCAD 86
```

```
Db 157 DTGFWCFT---MDPSVRWEYCNLTRCSD 181

RESULT 24
Q9UR5 PRELIMINARY; PRT; 113 AA.
AC Q9UR5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21181705; PubMed=11285247;
RA Ogorekova M., Kraft H.G., Ehnholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001)
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158663; AAF03680.1; -.
DR HSSP; P00747; 1PMK.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12685 MW; F3D65681D9B5253A CRC64;

Query Match 30.6%; Score 156; DB 4; Length 113;
Best Local Similarity 37.8%; Pred. No. 5.3e-11;
Matches 34; Conservative 12; Mismatches 32; Indels 12; Gaps 4;

QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATV--LQQTYYHAHRSDALQGLGKHNYCRNP 58
DQ : ||| ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
Db 9 RCYHGGQSGYRGFSFTVTGRTCSQSSSWTFRHQKTPENHPNDGLTM-----NYCRNP 63
QY 59 DNRERPWCYVQGLKPLV--QECMVHDCAD 86
DQ ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
Db 64 DADTGFWCFT---MDPSIRWEYCNLTRCSD 90

RESULT 25
O18783 PRELIMINARY; PRT; 806 AA.
AC O18783;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen.
OS Macropus eugenii (Tamar wallaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98004511; PubMed=9342350;
RA Lawn R.M., Schwartz K., Patthy L.;
RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog."
RL Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
```

```
DR EMBL; AF012297; AAB65760.1; -.
DR HSSP; P00747; 5HPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan.app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS50070; KRINGLE_2; 5.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 806 AA; 90981 MW; 95FAA86DC20064D5 CRC64;

Query Match 30.2%; Score 154; DB 6; Length 806;
Best Local Similarity 35.5%; Pred.No. 9e-10;
Matches 33; Conservative 11; Mismatches 25; Indels 24; Gaps 5;

QY 3 CYEENGHYRGKASTDTMGRPCLPWSNATVLQOTYHAHR-----SDALQLGLGKNY 54
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 371 CYEKGENYRGTSITISGKKQAWSS-----MTPGKHKTDPNFPNADLR-----NY 419

QY 55 CRNPENRRPWCYVOGLKPLV--QECMVHDC 85
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 420 CRNPDGKSPWCYT---MDPTVRWEFCNLKCS 449

RESULT 26
Q9UIK7 PRELIMINARY; PRT; 113 AA.
AC Q9UIK7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=21181705; PubMed=11285247;
RA Ogorelikova M., Kraft H.G., Ehrholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:1815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158659; AAF03678.1; -.
DR EMBL; AF158658; AAF03678.1; JOINED.
DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.

FT NON_TER 1
SQ SEQUENCE 113 AA; 12815 MW; 4F80ADF8708548CB CRC64;

Query Match 30.0%; Score 153; DB 4; Length 113;
Best Local Similarity 39.3%; Pred.No. 1.2e-10;
Matches 33; Conservative 8; Mismatches 35; Indels 8; Gaps 3;

QY 3 CYEENGHYRGKASTDTMGRPCLPWSNATVLQOTYHAHRSDALQLGLGKHNYCRNPENRR 62
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 11 CYHGDGQSYRGFSFTVTGRTCSWSSMT---PHWQRTTEYYPNGGLTRNYCRNPDAEI 67

QY 63 RPMCYYVQGLKPLV--QECMVHDC 84
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 68 RPMCYT---MDPSVRWEYCNLTRC 88

RESULT 27
Q8WNR1 PRELIMINARY; PRT; 359 AA.
AC Q8WNR1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
SEQUENCE FROM N.A.
RA Pirie-Shepherd S.R., Coffman K.T., Resnick D., Chan R., Kisker O.,
RA Folkman J., Waters D.J.;
RT "Angiostatin is Detectable in the Urine of Dogs with Spontaneous Bone
RT Cancer.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBAJ databases.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; AY069885; AAL58519.1; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003966; Prothrombin.
DR Pfam; PF00051; kringle; 4.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; PROTHROMBIN.
DR SMART; SM00130; KR; 4.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS50070; KRINGLE_2; 4.
KW Glycoprotein; Kringle.
FT NON_TER 1
SQ SEQUENCE 359 AA; 41172 MW; 776D35F4AB0BDD9E CRC64;

Query Match 29.7%; Score 151.5; DB 6; Length 359;
Best Local Similarity 44.3%; Pred.No. 6.6e-10;
Matches 31; Conservative 5; Mismatches 21; Indels 13; Gaps 3;

QY 3 CYEENGHYRGKASTDTMGRPCLPWSNATVLQOTYHAHRSDAL-----QLGLGKHNYCRN 57
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 282 CYHNGQSYRGTSITITGRKCSWSSMT-----PHRHKTPHFPEAGL-TWNYCRN 333

QY 58 PDNRRRPWCY 67
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 334 PDADKSPWCY 343

RESULT 28
Q9R0W3 PRELIMINARY; PRT; 812 AA.
AC Q9R0W3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen protein precursor (BC 3.4.21.7).
```

```
GN PLASMINOGEN.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bangert K., Johnsen A.H., Thorsen S.;
RT "Rat plasminogen: cDNA and gene structure.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91250378; PubMed=1645711;
RA Kanalas J.J., Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829 (1991).
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AJ242649; CAB46014.1; -.
DR HSSP; P00747; LPMK.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR001400; Somatotropin.
DR Pfam; PF00051; Kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR PRINTS; PRO1505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRY_PSP; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease; Signal.
FT SIGNAL
FT CHAIN
FT SEQUENCE 812 AA; 90535 MW; 8C703C51410EBC9E CRC64;
Query Match 29.3%; Score 149.5; DB 11; Length 812;
Best Local Similarity 37.8%; Pred. No. 2.8e-09;
Matches 34; Conservative 12; Mismatches 29; Indels 15; Gaps 6;
QY 3 CYEGNGHYRGKASTDIMGRCPLMNSATVLOQTYHAHRSALQL---GLGKHNYCRNP 59
Db 376 CYQNGKSYRGTSITNTCKKQSW-----VSMTPHSHSKTIPANFPDAGL-ENNYCRNP 429
QY 60 N-RRPWCYVQGLKPLV--QECMVHDCAD 86
Db 430 NDQRGWCFT---TDPVSRWEYCNLKRCS 456
RESULT 29
Q91691
ID Q91691 PRELIMINARY; PRT; 716 AA.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Growth factor Livertine.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RA Ruiz i Altaba A., Thery C.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; U57455; AAB52574.1; -.
DR HSSP; P00747; ICEA.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR PRINTS; PRO1505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRY_PSP; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 716 AA; 81971 MW; 508376A0E4398798 CRC64;
Query Match 29.2%; Score 149; DB 13; Length 716;
Best Local Similarity 33.0%; Pred. No. 2.8e-09;
Matches 33; Conservative 10; Mismatches 35; Indels 22; Gaps 3;
QY 2 TCYEGNGHYRGKASTDIMGRCPLMNSAT-----VLOQTYHAHRSALQLGLGKHNYCR 56
Db 289 TCFKRGEGYRGKANTTSGIFCQRWDSQTPOSHRFLPEKYPCKGLD-----ENYCR 340
QY 57 NPDNRRRPMCYVQV-----GLKPLVQECMVHDCADG 87
Db 341 NFDGSEAPWCFTTLFCMRMAYCFQIKRCKDDVLEPDCYHG 380
RESULT 30
Q8N2J4
ID Q8N2J4 PRELIMINARY; PRT; 381 AA.
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ90552.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovarian carcinoma;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 CUB DOMAIN.
DR EMBL; AK05033; BAC11365.1; -.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00431; CUB; 1.
```

